

Measurements of Solar Vector Magnetic Fields

*Proceedings of the MSFC workshop held at
George C. Marshall Space Flight Center
Huntsville, Alabama
May 15-18, 1984*



NASA Conference Publication 2374

Measurements of Solar Vector Magnetic Fields

Edited by
Mona J. Hagyard
George C. Marshall Space Flight Center

Proceedings of the MSFC workshop sponsored by the
National Aeronautics and Space Administration,
the University of Alabama in Huntsville, and the
National Oceanic and Atmospheric Administration
and held at George C. Marshall Space Flight Center
Huntsville, Alabama
May 15-18, 1984

NASA
National Aeronautics
and Space Administration
**Scientific and Technical
Information Branch**

1985

PREFACE

In recent years, new vector magnetographs and plans for future magnetograph systems seem to be springing up everywhere throughout the world-wide solar community, in the United States, France, China, Japan, the Soviet Union, and through the efforts of several international consortia. Though measurements of the photospheric vector field have been made since the late fifties and early sixties, primarily by our Russian colleagues, this recent revival of interest in developing vector magnetographs may be attributable to the advent of a modern technology that now allows us more versatility in approaching this formidable venture. In view of the rising interest in vector measurements, we here at the Marshall Space Flight Center believed it appropriate to convene a workshop dedicated to the subject of measurements of solar vector magnetic fields. Our primary goals were to enhance our understanding of the various techniques available for vector field measurements, to illuminate presently used methods for the interpretation of these measurements, and to develop dialogue between observers and theorists concerning the role of vector magnetic fields in our present understanding of solar processes.

The workshop was held at the Space Science Laboratory of the Marshall Space Flight Center, Huntsville, Alabama, U.S.A., in May of 1984. Twelve scientists from the international solar community joined over two dozen solar astronomers from the United States along with about thirty from the Huntsville area in vigorous and exciting discussions where new technical and theoretical ideas were exchanged and collaborative efforts initiated in an atmosphere of mutual cooperation.

The workshop was divided into four sessions, and each session began with invited talks followed by contributed papers and discussion periods. We thank the invited speakers for their excellent reviews that set the stage for the rest of the sessions. We are also extremely grateful for the high quality of all the contributed talks; they were a primary factor in the success of the workshop. The session chairmen were Yoshi Nakagawa, Vic Pizzo, Bob MacQueen, Eberhard Wiehr, Jack Harvey, Dave Rust and Einar Tandberg-Hanssen. We are indebted to these people for guiding the vigorous and sometimes intense discussions during their sessions.

We wish to acknowledge specially a few of the many people who made this workshop and its accompanying social program a reality. Financial support was provided by the following three institutions: the National Aeronautics and Space Administration through the Chief of the Office of Solar and Heliospheric Physics at NASA Headquarters, Dr. J.D. Bohlin; the University of Alabama in Huntsville, whose Director of Research, Dr. C. A. Lundquist, made possible the financial support of his institution and who arranged the Tuesday evening reception at the Burritt Museum; the National Oceanic and Atmospheric Administration, providing support through Gary Heckman, Chief of the Services Division of Space Environment Laboratory. We thank Dr. W. R. Lucas, Director of Marshall Space Flight Center, for providing resources and personnel of the Center for the workshop. In like manner, we thank Dr. Alex Dessler, Director of Space Science Laboratory, for

use of the facilities of the Laboratory; in addition, we are indebted to Alex whose idea it was in the first place to hold workshops like this one, and we appreciate his enthusiastic support of our efforts. To the members of the Space Science Laboratory's Solar Science Branch we extend our sincere thanks for their time and effort in helping us conduct the workshop and the extracurricular activities. Mrs. Tauna Moorehead provided invaluable assistance in preparing and editing the Proceedings, and we thank her for her time and effort. Last, but by no means least, we thank Mrs. Leila Reed, our secretary, for all that she did as we prepared for the workshop as well as for her efforts during the workshop; we are also indebted to her for her work in the preparation of these Proceedings.

Scientific Organizing Committee

Mona J. Hagyard, Chairman

Ernest Hildner

Ronald L. Moore

Steven T. Suess

Einar Tandberg-Hanssen

Shi Tsan Wu

TABLE OF CONTENTS

	Page
INTRODUCTION	1
SECTION 1. THEORETICAL UNDERSTANDING OF SOLAR VECTOR MAGNETIC FIELDS	5
The Vector Structure of Active Magnetic Fields. (Invited Paper) E. N. Parker	7
On the Numerical Computation of Nonlinear Force-Free Magnetic Fields S. T. Wu, et al.	17
Modeling Solar Magnetic Structures. (Invited Paper). B. C. Low	49
Magnetic and Thermodynamic Structure of the Solar Corona During Sunspot Minimum. V. A. Osherovich, et al.	66
Theoretical Approach to Evolution of Solar Magnetic Field Y. Nakagawa	78
Current Evolution in a Numerical Emerging-Magnetic-Flux Model T. G. Forbes	86
The Global Structure of Magnetic Fields Which Support Quiescent Prominences U. Anzer	101
SECTION 2. TECHNIQUES FOR MEASUREMENTS OF VECTOR FIELDS.	107
Trends in Measurement of Solar Vector Magnetic Fields Using the Zeeman Effect. (Invited Paper) J. W. Harvey	109
The Hanle Effect Applied to Magnetic Field Measurements. (Invited Paper). J. L. Leroy	121
Some Design Considerations for a Satellite-Borne Magnetograph D. M. Rust	141
Measurements of Vector Fields with a Diode Array E. J. Wiehr and W. Scholiers	153
Analysis of the New Polarimeter for the Marshall Space Flight Center Vector Magnetograph E. A. West	160

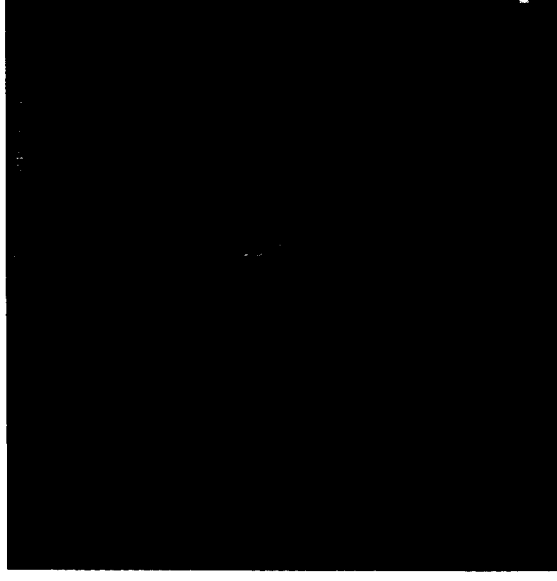
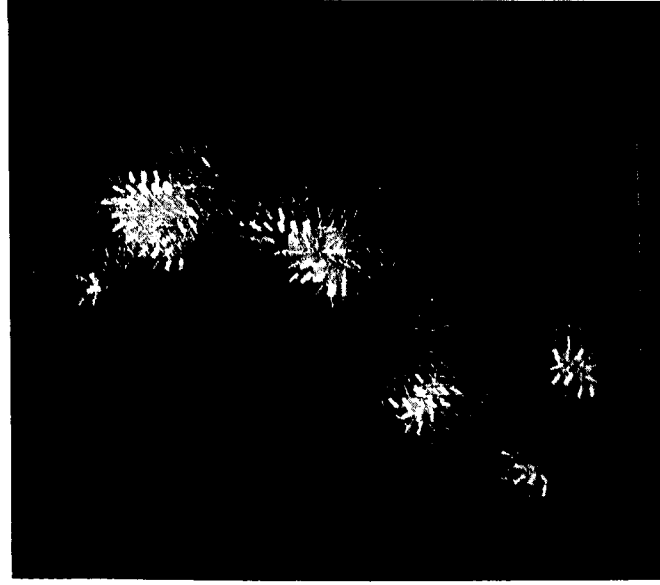
TABLE OF CONTENTS (Continued)

	Page
The Solar Vector Magnetograph of the Okayama Astrophysical Observatory M. Makita, et al.	173
Vector Magnetic Field Observations with the Haleakala Polarimeter D. L. Mickey	183
Stokes Parameters Modulator for Birefringent Filters A. Dollfus	192
The San Fernando Observatory Video Stokes Polarimeter P. H. Richter, et al.	202
The New Solar Magnetograph for the Canary Islands Observatory: T.H.E.M.I.S. J. Rayrole	219
The Vector Magnetograph of the Sayan Solar Observatory V. M. Grigoryev, et al.	231
Multichannel Birefringent Filter Ai Guoxiang and Hu Yuefeng	257
 SECTION 3. TECHNIQUES FOR INTERPRETATION OF OBSERVATIONAL DATA. . . .	 261
Diagnostics of Vector Magnetic Fields. (Invited Paper) J. O. Stenflo	263
Generation and Transfer of Polarized Radiation in the Solar Atmosphere: Physical Mechanisms and Magnetic-Field Diagnostics. (Invited Paper). E. Landi Degl'Innocenti	279
Dynamic Models of Flux Tubes in the Interpretation of Polarization Measurements. . E. Ribes, et al.	300
Least Squares Inversion of Stokes Profiles in the Presence of Velocity Gradients A. Skumanich, et al.	306
Contribution Functions for Zeeman-Split Lines and Line Formation in Photospheric Faculae A. van Ballegooijen	322
Effect of Polarized Radiative Transfer on the Hanle Magnetic Field Determination in Prominences: Analysis of Hydrogen H α Line Observations at the Pic-Du-Midi V. Bommier, et al.	335
The Effect of Line Damping, Magneto-optics and Parasitic Light on the Derivation of Sunspot Vector Magnetic Fields. (Abstract Only) A. Skumanich and B. W. Lites	341

TABLE OF CONTENTS (Concluded)

	Page
The Inference of Vector Magnetic Fields from Polarization Measurements With Limited Spectral Resolution Bruce W. Lites and Andrew Skumanich	342
Solar Monochromatic Images in Magneto-Sensitive Spectral Lines and Maps of Vector Magnetic Fields Ye Shihui, et al.	368
Magnetic Field Vector Measurements in Quiescent Prominences Via the Hanle Effect: Analysis of Prominences Observed at the Pic-Du-Midi and at Sacramento Peak. V. Bommier, et al.	375
Inferred Flows of Electric Currents in Solar Active Regions Y. J. Ding, et al.	379
Comparisons of Simultaneous Vector Magnetograms M. Makita, et al.	399
Ephemeral Regions Vs. Pseudo-Ephemeral Regions. S. F. Martin, et al.	403
Evidence for Submergence of Magnetic Flux in a Growing Active Region. (Abstract Only) D. M. Rabin, et al.	437
SECTION 4. TECHNIQUES FOR DATA DISPLAY	439
Solar Physics Applications of Computer Graphics and Image Processing. M. D. Altschuler	441
Graphic Displays of Vector Magnetograph Data D. M. Rabin and E. A. West	454
LIST OF PARTICIPANTS	467

Measurements of Solar Vector Magnetic Fields



Vector magnetic field data from the Okayama Astrophysical Observatory (courtesy of M. Makita). The observations are of Active Region 4474 on April 27, 1984. This region produced numerous flares during its disk passage, the largest being an X13/3B on April 24.

INTRODUCTION

There is no question today among solar astronomers about the fundamental importance of the solar magnetic field in most of the physics of the Sun. In the visible layers of the solar atmosphere we see this field shaping and controlling the solar plasma, providing the energy source for the explosive phenomenon of solar flares, forming sunspots, appearing in the supergranular cells as subarcsecond concentrations of kilogauss field strengths, and interacting with the convective motions of the subsurface plasma to produce the solar cycle. This fundamental role of the magnetic field has prompted the development over the years of instruments to measure the field, beginning with the pioneering work of George Ellery Hale in the early years of this century. In the eight decades since Hale's initial work in measuring the magnetic fields in sunspots, we have progressed steadily in developing magnetographs, and so-called "line-of-sight" or "longitudinal" magnetographs to measure the line-of-sight component of the magnetic field in the photosphere are common instruments at many solar observatories today. However, to fully understand the solar field, we must be able to measure the complete magnetic vector, and this has always proven to be a more difficult achievement. The measurement of the transverse (to the line-of-sight) magnetic component requires sensitive detectors, precision polarimetry, and minimal or minute instrumental polarization in the optical system. In addition to the instrumental problems associated with vector magnetographs, there is the perplexing issue of interpreting the data once they are obtained. The theory of the radiative transfer of polarized light in a magnetic medium does not provide easily managed formulas to relate measured values to magnetic field components except under sometimes questionable assumptions.

With recent developments in instrumental technology and the availability of high-powered computers, we are now in a better position to develop vector magnetograph systems that will reliably measure the vector field and easily manipulate the data for more efficient analyses. With many observatories taking advantage of this new technology by upgrading existing magnetographs or building new ones, it seemed that it would be useful to provide a forum for exchanging new ideas and concepts among those involved in developing the instrumentation and the interpretative methods. The highly successful MSFC Workshop was such a forum, and consequently, this volume of proceedings from that workshop provides today's solar astronomer with an overview of current instrumental, analytical, observational and theoretical research in the area of solar vector magnetic fields.

In the first section of these Proceedings, our theoretical understanding of solar vector magnetic fields is explored. Our objective in this session of the workshop was to generate dialogue between theorists and observers to determine what aspects of solar vector fields should be investigated. Questions we asked were: what aspects of vector magnetic fields are important to theoretical models? Where should the field be measured? In what solar processes and phenomena are they important? What are the important derived quantities? In this section, E. N. Parker points out that one of the immediate rewards of measuring the vector field at the photospheric surface is information on the strains introduced into the magnetic field by the convective motions that deform the field. However, Parker also indicates that the rapid shuffling of the magnetic field lines by granular motions produces small-scale fields that are not in static equilibrium, so that we cannot extrapolate our vector measurements of the photospheric field upwards with equations

assuming static equilibrium. In his paper in this session, B. C. Low attempts to assess what useful physics can be learned using vector magnetograph data as inputs to model quasi-static, large-scale magnetic structures. Low shows that the existence of force-free states can be tested using vector magnetic field data in equations derived from classical virial relations. In regions where the field is force-free, Low points out that the vector data can determine whether it is of the linear or non-linear type. Low also reports on recent progress in developing theoretical models of three-dimensional plasma loops.

In Section 2, techniques for measurements of vector magnetic fields are presented. This session was devoted primarily to the instrumentation required for measuring the vector field in the photosphere, transition region, and corona. Both ground-based and flight systems were included, and the purview included systems in operation, in development, and in the conceptual phase. J. W. Harvey introduces the subject by reviewing recent trends in spectropolarimetry as applied to measurements of the Zeeman effect. He discusses recent developments in detector arrays that improve observing efficiency and the use of narrow-band filters and Fourier transform spectrometers instead of grating spectrographs to improve angular and temporal coverage on the one hand and to improve spectral coverage and precision on the other. In the second review paper in this section, J. L. Leroy gives a detailed exposition on the application of the Hanle effect to the measurement of weak magnetic fields. Leroy reviews the successful development recently of the exact quantum-mechanical treatment of this effect, and he indicates the various astrophysical applications of the technique, for example, measurements of prominence fields and the turbulent photospheric field.

Section 3 is devoted to techniques for interpretation of the observational data and to some observational results. This session encompassed such topics as magneto-optical effects, anomalous Stokes profiles, unresolved fine structure, the Hanle effect, line damping, and inversion techniques; several papers on observational results were also presented. J. O. Stenflo, in his paper "Diagnostics of Vector Magnetic Fields," discusses the fundamental problem of interpretation that is introduced by the spatially unresolved fine structure of the magnetic field at the photospheric level. Other diagnostic problems of fluxtube physics are discussed, including mass motions inside fluxtubes, temperature structure as a function of fluxtube diameter, and the divergence of the fluxtube field with height. Stenflo indicates that the resolution of this diagnostic problem lies in either observations in suitable combinations of spectral lines or in deriving the Stokes profiles with very high spectral resolution (as provided, for example, by the Fourier Transform Spectrometer of the NSO McMath telescope). In his review paper, E. Landi Degl'Innocenti discusses first the two different mechanisms that produce polarization in solar spectral lines: the Zeeman effect and atomic polarization. He then turns to the inverse problem, deducing the value of the magnetic field from polarimetric measurements, and reviews the factors that make this problem basically still unresolved. Landi gives an overview of the different algorithms that have been developed to deduce the magnetic field from profiles of the Stokes parameters and points out the reasons why the different algorithms may produce different results for the measurement of the magnetic field.

In this age of advanced computer technology, modern techniques to display computer-generated graphics can provide sophisticated methods for visualization of the observed vector field and its derivative parameters as they evolve and change with time. It was the purpose of the final

session of the workshop to explore some of these techniques that have been or can be developed. To provide an overview of the subject, we asked M. Altschuler of the University of Pennsylvania School of Medicine to "return to the fold" of solar physics briefly and describe to us what can be done in the field of computer graphics and image processing today. In his review paper, Altschuler describes what is meant by computer graphics and computer imaging, and then discusses how they can be used by solar physicists to process raw images, to compare different data, to transform data to different coordinate systems, to create data bases from processed or unprocessed data, and to interrogate and visualize the information in these data bases. He then outlines some important considerations to keep in mind in setting up a graphics and image-processing facility.

At the completion of this workshop there were admittedly many unresolved problems and unanswered questions. However, those of us who participated in the workshop were caught up in an atmosphere of renewed interest in measuring the Sun's vector magnetic field, generated by the exchange of new ideas and techniques that did occur. It is our hope and expectation that these new concepts will guide and stimulate further work in this field, both theoretical and technical, leading us to new breakthroughs in addressing the problems that remain. We also anticipate that these Proceedings will be useful to the Workshop participants and to the broader solar research community.

SECTION 1

THEORETICAL UNDERSTANDING OF SOLAR VECTOR MAGNETIC FIELDS

THE VECTOR STRUCTURE OF ACTIVE MAGNETIC FIELDS

E.N. PARKER

Dept. of Physics, University of Chicago, Chicago, Illinois

Abstract. Observations are needed to show the form of the strains introduced into the fields above the surface of the sun. The longitudinal component alone does not provide the basic information, so that it has been necessary in the past to use the filamentary structure observed in H_α to supplement the longitudinal information. Vector measurements provide the additional essential information to determine the strains, with the filamentary structure available as a check for consistency. It is to be expected, then, that vector measurements will permit a direct mapping of the strains imposed on the magnetic fields of active regions. It will be interesting to study the relation of those strains to the emergence of magnetic flux, flares, eruptive prominences, etc. In particular we may hope to study the relaxation of the strains via the dynamical nonequilibrium.

I. DETERMINATION OF MAGNETIC FIELDS

The vector nature of magnetic fields is well known. The purpose of this review is to outline the connection between the topology of the vector field and its equilibrium properties. As we shall see, the topology determines the equilibrium configuration, and even whether there is an equilibrium. For as we shall see, some topologies cannot be in equilibrium, forming instead sheet-like discontinuities where the fluid is in dynamical motion and the field subject to rapid dissipation.

The development of the subject makes it clear that an observational determination of the complete vector field at the photosphere is essential to deducing the configuration of the field above the surface of the sun. For the fact is that we are able to observe the magnetic field of the sun only in a thin layer at the photosphere. Some variation in height ($\sim 2 \times 10^3$ km) is possible through the choice of different spectral lines, but essentially all observational information is confined to a sheet. The problem is to deduce the rest of the field from the complete vector $\mathbf{B}(x, y, 0, t)$ at that one sheet.

Extrapolation downward from the surface is a dubious exercise, even in principle, because of the convection $\mathbf{v}(\mathbf{r}, t)$ from the surface of the dense gas, overpowering and deforming \mathbf{B} in a manner described by the familiar induction equation

$$\partial \mathbf{B} / \partial t = \nabla \times (\mathbf{v} \times \mathbf{B}) + \eta \nabla^2 \mathbf{B} \quad (1)$$

where the resistive decay coefficient $\eta = c^2 / 4\pi\sigma$ is negligible ($< 10^9$ cm²/sec) for most considerations. Even in the intense magnetic fields of sunspots, which are so strong as to suppress the convective motion \mathbf{v} in large degree, the equations for static equilibrium in a gravitational potential ϕ and fluid pressure p ,

$$\nabla (p + \mathbf{B}^2 / 8\pi) = (\mathbf{B} \cdot \nabla) \mathbf{B} / 4\pi - \rho \nabla \phi \quad (2)$$

tell us only that the interior of the field is strongly evacuated, so as to compress the field to the observed high intensity of 3000 gauss. More cannot be said

until a knowledge of the fluid pressure is available. Unfortunately the vertical structure of a sunspot extends downward over many scale heights, so that ρ is highly sensitive to the temperature T , which is highly sensitive in turn to whatever convective velocity \mathbf{v} may exist in and around the intense magnetic field. As a result, knowledge of the basic causes of sunspots and of their subsurface structure extends little if at all beyond conjecture (see discussion in Meyer, Schmidt, Weiss, and Wilson, 1974; Parker, 1979a).

Above the surface of the sun the situation is somewhat different because above active regions the fluid pressure declines rapidly to values well below the magnetic stress level $B^2/8\pi$. What is more, we can see the filamentary structure of the gas embedded in the field, providing direct information on the topology of the field. So we are still faced with the basic magnetohydrodynamical relations (1) and (2), but we have supplementary information to assist us on our way.

There are many dynamical phenomena, e.g. spicules, coronal transients, flares, the solar wind, etc. arising in the magnetic fields above the surface of the sun. They involve the dynamical equation

$$\rho \left[\partial \mathbf{v} / \partial t + (\mathbf{v} \cdot \nabla) \mathbf{v} \right] = -\nabla(\rho + B^2/8\pi) + (\mathbf{B} \cdot \nabla) \mathbf{B} / 4\pi - \rho \nabla \phi \quad (3)$$

in place of the static equation (2). To keep the present exposition as simple as possible we avoid such effects as much as possible. Some of the other lectures will deal directly with them. We pursue, instead, the "simple" concept of the quasi-static fields above the surface of the sun and what we can learn about them from the theoretical properties of the static equation (2).

We are all familiar with static electric and magnetic fields in vacuo. Each can be described in terms of a single scale function satisfying Laplace's equation in any region free of separated electric charges or free of electric currents, respectively. Such fields have the convenient property that they are uniquely determined throughout any volume of space by the specification of the normal component over the surface enclosing that volume. Thus, for instance, the line-of-sight component of the magnetic field on the surface of the sun, together with the knowledge that the field falls off to negligible values at large distances above the surface, is sufficient to determine the complete static magnetic field above the surface of the sun. In fact, the magnetic field is nothing more than the field configuration with minimum energy for the given normal component on the boundaries. The topology of the field is a secondary consideration, taking on the simplest form compatible with the boundary conditions (see Adams and Pneuman, 1976; Sheeley, 1982, for examples).

It is an interesting observational fact, first pointed out by Sheeley et al. (1975) that the large-scale magnetic field above an emerging active region quickly approaches a simple potential form, with the most direct connection and the minimum energy, for the distribution of the normal component of the field observed at the surface. It is well known, however, that the magnetic fields at the surface of the sun are continually deformed by the motions of the dense fluid at and below the surface. Such convective motions continually deform the field, introducing strains that spread themselves along the entire length of the lines of force above the surface. The result is that the field above the active regions is not a purely potential form on small scales at least, but has a more complicated form high in the atmosphere where $\rho \ll B^2/8\pi$. This has led to the study of equation (2) with small or negligible gas pressure, so that

$$(\nabla \times \mathbf{B}) \times \mathbf{B} = 0 \quad (4)$$

for the "force-free" magnetic field, i.e. for the field which exerts no force on the gas throughout the solar atmosphere above the visible surface. The general solution of (4) is

$$\nabla \times \mathbf{B} = \alpha(\mathbf{r}) \mathbf{B} \quad (5)$$

where $\alpha(\mathbf{r})$ is a function of position that is arbitrary except for being constant along each magnetic line of force, $\mathbf{B} \cdot \nabla \alpha = 0$. It is at this point that the topology of the magnetic field makes its appearance, because the field lines represent lines of constant α which must be determined before \mathbf{B} can be computed and related to the observed field at the surface of the sun. This problem—which is sufficiently difficult that it may fairly be called a dilemma—can be avoided by expanding $\alpha(\mathbf{r})$ about some point \mathbf{r}_0 in the field and keeping only the first term in the series, $\alpha(\mathbf{r}_0) = \text{constant}$. The topology of the field drops out of the picture, then, because $\nabla \alpha = 0$ and $\mathbf{B} \cdot \nabla \alpha = 0$ for any field topology. As a result, a variety of models of force-free fields with constant α have been worked out for comparison with observation. Such models have enjoyed a measure of success as the next step beyond the potential solutions in which α is equated to zero (cf. Raadu and Nakagawa, 1971; Nakagawa, Raadu, Billings, and McNamara, 1971; Nakagawa and Raadu, 1972; Tanaka and Nakagawa, 1973; Molodensky, 1974, 1975, 1976; Sheeley and Harvey, 1975; Levine, 1975; Low and Nakagawa, 1975; Low, 1977, 1978; Migliuolo and Cargill, 1983).

It is at this point in the theory of static equilibrium of magnetic fields in a highly conducting medium that we encounter the necessity of vector field measurements. If $\alpha = 0$, then the solution of (5) is uniquely determined by the normal component (often the line of sight component) over the surface of the sun. But if $\alpha \neq 0$, the precise value of α has to be determined, even assuming that it is uniform over the region. The line of sight component is sufficient to determine the solution only if α is already known. There is, of course, other information available, such as the filamentary structure of the gas observed in the photosphere in H_α and other lines, which shows the local direction of the magnetic vector even if it does not provide a direct quantitative measurement. The filaments have been used to supply the value of α , allowing a unique determination of the field from the observed normal component (cf. Nakagawa, Raadu, and Harvey, 1973; Sheeley, 1982). The theoretical studies of force-free field with constant α show, too, that the observations are not fitted precisely by this special class of field. Indeed, the deviations may be large on both large and small scales in many cases. And it is at this point that quantitative measurement of both the longitudinal (line of sight) and transverse components of the magnetic field becomes essential. The observed filamentary structure at the surface of the sun is not sufficient.

To be specific, suppose that the observer examines the central region of the solar disk and observes the line of sight component and the two transverse components of the magnetic field. The line of sight component we designate by B_z , using a coordinate system in which the z-axis is perpendicular to the local solar surface and the x and y axes lie in the surface. The two transverse components $B_x(x, y)$ and $B_y(x, y)$ allow the z-component of the curl, $\partial B_y / \partial x - \partial B_x / \partial y$, to be computed, which determines α as a function of position upon substitution

into (5). This exercise provides a direct measure of the torsion of $\alpha(x, y)$ and an illustration of the degree of error in the assumption that α is a constant. It illustrates, too, the essential role of the transverse components of the magnetic field, without which one can do little to determine the manner of variation of α .

Later papers in this workshop go into the problem of computing force-free fields and determining α . With the assumption that α is uniform, the problem is a linear one, as is obvious by inspection of equation (5). As soon as one admits a variable α the problem becomes nonlinear in a particularly ugly way. The basic problem is that (5) is a linear equation in \mathbf{B} only if α is known independently, not only on the photosphere where it may be derived directly from observation, but throughout the entire volume above the photosphere where the field is non-negligible. Since $\mathbf{B} \cdot \nabla \alpha = 0$, α projects upward from the photosphere along the field that we have not yet calculated. Thus, the calculation of the field is tied directly to the final configuration of the field. Now in a magnetically isolated (closed) region on the surface of the sun, any line of force that extends up from a point P_1 must re-enter the sun at some other point P_2 in the same region where the value of α is the same as α at P_1 . So if we plot contours of constant α on the surface of the sun, as sketched in Figure 1, we know that any line of force of a force-free field extending up from a point on a given contour $\alpha = \alpha_0$ must re-enter the sun somewhere on the same contour, or on a nearby contour for the same value of α_0 . But where on the contour is not clear. That depends upon the net torsion introduced into the field by the fluid motions below the surface. The function α is itself a measure of the torsion along its own line of force. To illustrate this directly, consider a small area S oriented perpendicular to the field \mathbf{B} at some point P in space. Then if C is the closed curve around the boundary of S , the integral of (5) over S yields

$$\int d\mathbf{S} \cdot \nabla \times \mathbf{B} = \alpha \int d\mathbf{S} \cdot \mathbf{B}.$$

The integral on the right hand side of the equation is just the total flux Φ across S , while the integral on the left side can be transformed by Stokes theorem into a line integral of the transverse component of the field around the periphery,

$$\oint_C d\mathbf{s} \cdot \mathbf{B} = \alpha \Phi \quad (6)$$

To state this relation in its most primitive form, consider a circular contour with small radius r . Then $\Phi = \pi r^2 B$, where B is the magnitude of \mathbf{B} at the center of S . Denote by B_\perp the mean value of the transverse field circling C

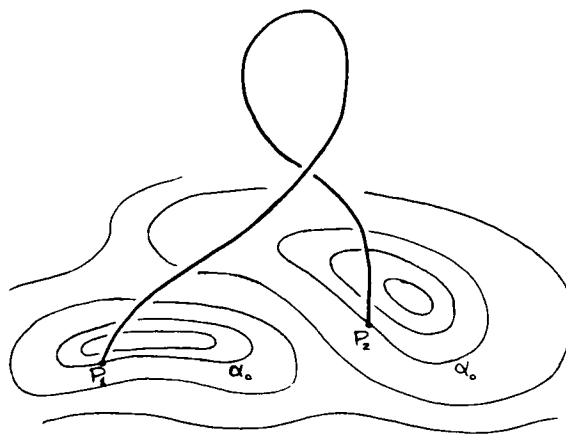


Fig. 1. A sketch of the contours of constant α in the photosphere, with the heavy looping line representing a magnetic line of force of the force-free field with an assigned value α_0 connecting between two points P_1 and P_2 .

so that the left hand side of (6) can be written $2\pi r b_\varphi$. Then (6) reduces to

$$b_\varphi = \frac{1}{2} \alpha r B \quad (7)$$

showing that the torsional component of the field, arising from the overall twisting α , grows linear with radius r at a rate proportional to α . It is evident, then, that the function $\alpha(r)$ provides the information on the torsion, telling us how far around the contours for constant α the fields are twisted. But it is a difficult nonlinear mathematical problem to provide a precise result for α and the associated $B(r)$.

The purpose here is to pursue the theoretical problem beyond these first stages to provide a broader view of the implications of static equilibrium of a magnetic field, described by (2), with particular attention to field topologies for which there is no equilibrium at all.

II. INDETERMINATE MAGNETIC FIELDS

The general theoretical question posed by solution of the static equilibrium equation (2) for the magnetic fields arching above the surface of the sun does not depend critically on the curvature of the field. It is essentially the problem of a magnetic field B extending in the positive z -direction between two conducting planes $z=0, h$. The field is anchored in the two planes, which represent the dense gases of the photosphere. The effects of gravity do not alter the essential nature of the basic problem, so we consider solutions to

$$\nabla(p + B^2/8\pi) = (B \cdot \nabla)B/4\pi \quad (8)$$

in the region $0 < z < h$, with the fluid pressure p controlled at the bounding surfaces $z=0, h$, just as p is determined at the photosphere of the sun. We suppose that an observer is able to supply us with the magnetic field

$B(x, y, 0)$ and $B(x, y, h)$ at $z=0, h$. And then we play a more intimate game, involving direct manipulation of the system by the convective motions, introducing various displacements of the field at $z=0$, while holding things fixed at $z=h$. We may imagine that the fluid motions at $z=0$ shuffle the footpoints of the lines of force through various disordered arbitrary patterns, as sketched in Figure 2. The shuffling of the footpoints produces a wrapping of the lines of force about each other that spreads out more or less uniformly along the lines of force all the way to the other end at $z=h$.

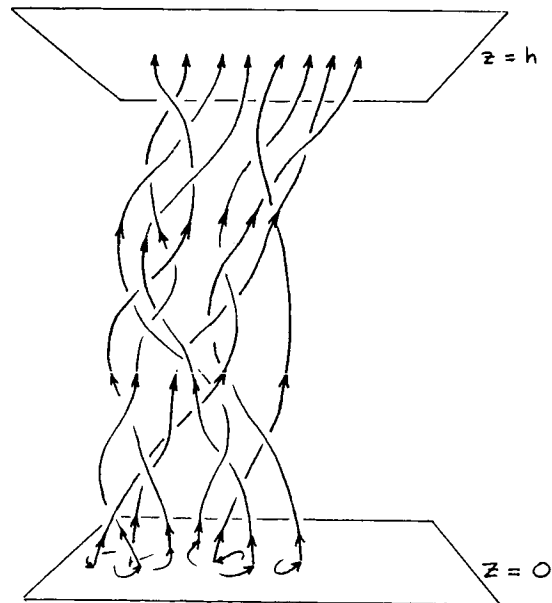


Fig. 2. A sketch of magnetic lines of force extending from $z=0$ to $z=h$, with their footpoints shuffled at $z=0$ in some irregular and changing pattern, with the result that the lines of force are wrapped and woven about each other in irregular topological patterns.

We might wonder, then, how the terrestrial observer should go about deducing the accumulated wrapping of the field from the transverse components of the field observed at any instant of time at $z=0, h$. The observer cannot follow the shuffling process, of course, because the motion of the footpoints at the photosphere is not readily detectable, and, in any case, there was never an opportunity to start from zero wrapping. So how can the wrapping, i.e. the present field configuration $\mathbf{B}(x, y, z)$, be worked out throughout $0 < z < h$, from the static equilibrium equation (2) and the fields at $z=0, h$?

The answer appears to be that the field cannot be deduced, in principle, for the simple reason that the random shuffling of the footpoints of the field at zero generally produces field topologies for which there is no solution to equation (2), for any fluid pressure applied at $z=0, h$.

The absence of an equilibrium may be seen from the fact that, in the present situation, without gravity, the pressure satisfies $\mathbf{B} \cdot \nabla p = 0$, i.e. the fluid pressure p is constant along each line of force. It follows that $p(x, y, h)$ must match to $p(x, y, 0)$ projected along the lines of force to $z=h$. This condition is not generally satisfied in the sun, nor is it generally so strongly violated (except in sunspots) as to be of great interest. So suppose that $p(x, y, z)$ is properly coordinated with the topology of the field and $p(x, y, h)$ is consistent with $\mathbf{B} \cdot \nabla p = 0$.

The first point is that there are no solutions to the equilibrium equation (2) throughout $0 < z < h$ in which the wrapping pattern varies along the general direction of the field. That is to say, there is no equilibrium of the field throughout the region if the extending shuffling of the footpoints of the lines of force is carried through more than one pattern, for the pattern of the shuffling becomes the pattern of the wrapping (Parker, 1972, 1979b, pp. 359-391; 1981a,b; Low, 1980, 1981, 1982; Tsinganos, Distler, and Rosner, 1984; Yu, 1973).

We may be sure that the disorderly convection in sun introduces highly variable shuffling patterns, so that the wrapping of the lines is generally irregular. The nonequilibrium result is the production of thin current sheets, at which neutral point reconnection cuts across the lines of force and continually reduces the topology toward a pattern of uniform wrapping.

Now if we suppose that the extended shuffling of the footpoints carries each footpoint around only within some local domain with dimensions l small compared to the length h of the lines of force, it follows that there are few, if any, lines of force extending large transverse distances across the field. The wrapping patterns are local, with a scale l , and when they have been reduced by neutral point reconnection to uniformity along the z -direction we have $\partial B_z / \partial z = 0$ through the broad interior of $0 < z < h$. Hence $\partial B_x / \partial x + \partial B_y / \partial y = 0$ and there are no sources or sinks of the transverse field (B_x, B_y) . Hence the transverse field forms locally closed patterns, as sketched in Fig. 3. That is to say, the total field is now equivalent to a bundle of close packed individual twisted flux tubes, sketched in Fig. 4. Equation (2) reduces to the familiar Grad-Shafranov equation for

$$B_x = +\partial A / \partial y, \quad B_y = -\partial A / \partial x, \quad B_z = B_z(A), \quad p = p(A) \quad (9)$$

$$\nabla^2 A + 4\pi p'(A) = 0, \quad p(A) \equiv p(A) + B_z^2(A) / 8\pi, \quad (10)$$

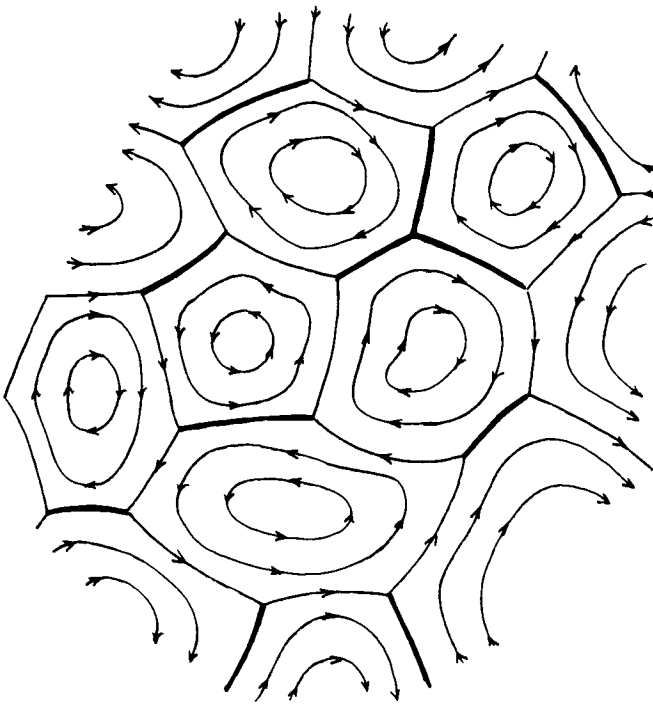


Fig. 3. A sketch of the transverse field patterns on a surface $z=\text{constant}$. The heavy lines represent neutral (current) sheets where opposite fields press together.

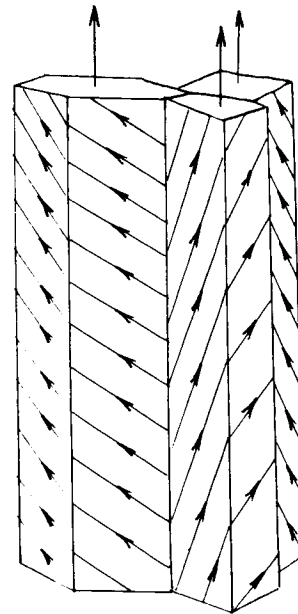


Fig. 4. A sketch of an assortment of twisted flux tubes packed together to form a continuum field.

where $P(A)$ is an arbitrary function of its argument. It is readily shown, however, that a close-packed bundle of twisted flux tubes does not provide a solution to (9) and (10) (Parker, 1983a,b,c,d) because when packed together the individual flux tubes generally meet at three-fold vertices, as sketched in Fig. 3. A four-fold vertex requires that two three-fold vertices be pressed together with precisely the right force to meet at a common point, as sketched in Fig. 5a,b. Too much or too little force leaves two three-fold vertices, sketched in Fig. 5c,d. An assortment of twisted flux tubes of random diameters and torsions does not have this property. A three-fold vertex cannot be represented by a single-valued continuous solution to (10). Equilibrium is not possible with three-fold vertices because the transverse fields of at least two of the three twisted tubes that form each vertex must reverse sign across the boundary between the two tubes. The result is neutral point reconnection, coalescing the transverse fields of the tubes with the same sense of twist, as sketched in Fig. 6a,b. The phenomenon is known as the "coalescence of islands" in laboratory magnetized plasmas. The nonequilibrium reconnection goes on until the entire bundle of twisted tubes is reduced to two regions of opposite twist (Parker, 1982,1983a,c,d,e).

It would appear that much, if not all, of the heating of the active corona of the sun, and other stars, is accomplished in this way (Parker, 1983b) with the work done (by the photospheric convection in shuffling the field) going into heat through the mechanism of neutral point reconnection. Indeed, the entire nonequilibrium process among close packed twisted flux tubes is equivalent to two dimensional magnetohydro-

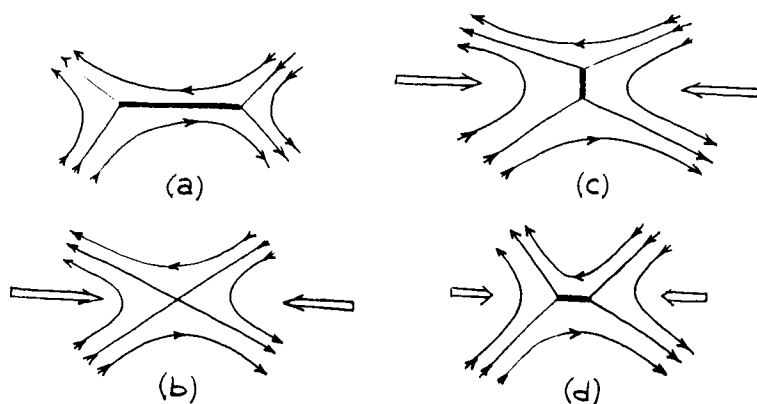


Fig. 5.(a) A sketch of two three-fold vertices in the transverse field pattern, where four twisted flux tubes meet. The heavy line indicates the neutral sheet. (b) Forces indicated by the two broad arrows press the two opposing twisted tubes together with precisely the right force to coalesce the two three-fold vertices into one four-fold vertex. (c) A force that is slightly too strong or (d) slightly insufficient leaves two three-fold vertices.

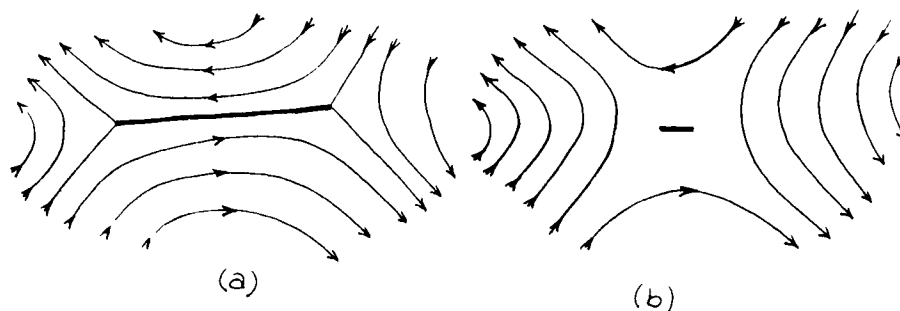


Fig. 6.(a) A sketch of the magnetic fields around three-fold vertices illustrating the unavoidable fact that at least two of the three fields around any such vertex must be oppositely directed. This leads to neutral point reconnection between the two opposite fields, with the coalescence of the fields of two of the cells illustrated in (b). The heavy lines indicate neutral sheets.

dynamic turbulence, whose internal dynamical nature has been studied quantitatively by Fyfe and Montgomery (1976); Fyfe, Joyce, and Montgomery (1977); Fyfe, Montgomery and Joyce (1977); Montgomery and Vahala (1979); Matthaeus and Montgomery (1981).

The implications for the present discussion of the determination of fields above the surface of the sun are simply that the problem must be approached on both a large and a small scale. The rapid shuffling of the lines of force by the granules suggests that the small-scale fields are never near equilibrium, so we cannot deduce them from the fields observed at the surface of the sun by extrapolating upward with the equation for static equilibrium. On a large-scale the fields may approach static equilibrium. But the large-scale field is not composed, then, of a continuum of elemental flux tubes whose tension along the large-scale field is twice the isotropic

pressure of the field. Instead, the large-scale field is made up of twisted elemental flux tubes in which the mean tension is different from twice the mean pressure exerted on their neighbors. Hence the conventional static equilibrium equation (2) does not apply precisely, and might lead to error if applied to complicated large-scale configurations. Specifically, we have estimated (Parker, 1983b) (but cannot demonstrate) that the transverse component of the field may be of the order of one quarter of the longitudinal component. Since the alteration of the tension and pressure depend upon the square of the transverse field, the effect is of the order of one part in sixteen, and would probably introduce no great error in calculating fields with simple large-scale configurations. Ultimately, of course, we need the transverse components of the field with high spatial resolution, from an instrument with the quality of the projected Solar Optical Telescope. For we must get the precise picture of the small-scale fields before really complicated large-scale field configurations can be calculated.

So that is where we seem to be with regard to reconstructing the magnetic fields above the surface of the sun from the vector fields observed at the surface. The reconstruction will be an interesting task. As noted above, there are several papers to be presented at this workshop which present specific examples of the reconstruction procedure.

While we are engaged in measuring and plotting vector fields at the surface of the sun, it is clear that the immediate rewards will be a wealth of information on the strains in the magnetic fields in active regions, through various stages of activity, with many curious results, we can be sure. We will perhaps see better where and when magnetic field is retracted back into the sun, what magnetic strains are associated with coronal transients and with flares, what strains show up in newly emerging active regions, and what, if any, strain patterns appear in association with the formation and dissolution of sunspots.

References

- Adams, A. and Pneuman, G.W.: 1976, Solar Phys. 46, 185.
 Fyfe, D., Joyce, G., and Montgomery, D.: 1977, J. Plasma Phys. 17, 317.
 Fyfe, D. and Montgomery, D.: 1976, J. Plasma Phys. 16, 181.
 Fyfe, D., Montgomery, D., and Joyce, G.: 1977, J. Plasma Phys. 17, 369.
 Levine, R.H.: 1975, Solar Phys. 44, 365.
 Low, B.C.: 1977, Astrophys. J. 212, 234.
 Low, B.C.: 1978, Astrophys. J. 224, 668.
 Low, B.C.: 1980, Astrophys. J. 239, 377.
 Low, B.C.: 1981, Astrophys. J. 251, 352.
 Low, B.C.: 1982, Astrophys. J. 263, 952.
 Low, B.C. and Nakagawa, Y.: 1975, Astrophys. J. 199, 237.
 Matthaeus, W.H. and Montgomery, D.: 1981, J. Plasma Phys. 25, 11.
 Meyer, F., Schmidt, H.U., Weiss, N.O., and Wilson, P.R.: 1974, Mon. Not. Roy Astron. Soc. 169, 35.
 Migliuolo, S. and Cargill, P.J.: 1983, Astrophys. J. 271, 820.
 Molodensky, M.M.: 1974, Solar Phys. 39, 393.
 Molodensky, M.M.: 1975, Solar Phys. 43, 311.
 Molodensky, M.M.: 1976, Solar Phys. 49, 279.
 Montgomery, D. and Vahala, G.: 1979, J. Plasma Phys. 21, 71.
 Nakagawa, Y. and Raadu, M.A.: 1972, Solar Phys. 25, 127.

- Nakagawa, Y., Raadu, M.A., Billings, D.E., and McNamara, D.: 1971, Solar Phys. 19, 72.
- Nakagawa, Y., Raadu, M.A., and Harvey, J.W.: 1973, Solar Phys. 30, 421.
- Parker, E.N.: 1972, Astrophys. J. 174, 499.
- Parker, E.N.: 1979a, Astrophys. J. 230, 905.
- Parker, E.N.: 1979b, Cosmical Magnetic Fields, Oxford, Clarendon Press.
- Parker, E.N.: 1981a, Astrophys. J. 244, 631.
- Parker, E.N.: 1981b, Astrophys. J. 244, 644.
- Parker, E.N.: 1982, Geophys. Astrophys. Fluid Dyn. 22, 195.
- Parker, E.N.: 1983a, Astrophys. J. 264, 635.
- Parker, E.N.: 1983b, Astrophys. J. 264, 642.
- Parker, E.N.: 1983c, Geophys. Astrophys. Fluid Dyn. 23, 85.
- Parker, E.N.: 1983d, Geophys. Astrophys. Fluid Dyn. 24, 79.
- Parker, E.N.: 1983e, Geophys. Astrophys. Fluid Dyn. 24, 245.
- Raadu, M.A. and Nakagawa, Y.: 1971, Solar Phys. 20, 64.
- Sheeley, N.R.: 1982, Astrophys. J. 255, 316.
- Sheeley, N.R., Bohlin, J.D., Brueckner, G.E., Purcell, J.D., Scherrer, V., and Tousey, R.: 1975, Solar Phys. 40, 103.
- Sheeley, N.R. and Harvey, J.W.: 1975, Solar Phys. 45, 275.
- Tanaka, K. and Nakagawa, Y.: 1973, Solar Phys. 33, 187.
- Tsinganos, K.C., Distler, J., and Rosner, R.: 1984, Astrophys. J. 278, 409.
- Yu, G.: 1973, Astrophys. J. 181, 1003.

ON THE NUMERICAL COMPUTATION OF NONLINEAR FORCE-FREE MAGNETIC FIELDS

S. T. Wu and H. M. Chang

The University of Alabama in Huntsville
Huntsville, Alabama 35899, USA

and

M. J. Hagyard

NASA/Marshall Space Flight Center
Huntsville, Alabama 35812, USA

ABSTRACT

An algorithm has been developed to extrapolate nonlinear force-free magnetic fields from a source surface, given the proper boundary conditions. In this paper we present the results of this work; describing the mathematical formalism that was developed, the numerical techniques employed, and the stability criteria developed for these numerical schemes. An analytical solution is used for a test case; the results show that the computational accuracy for the case of a nonlinear force-free magnetic field was on the order of a few percent ($<5\%$).

1. Introduction

Observations have shown that physical conditions in the solar atmosphere are strongly controlled by solar magnetic fields. The appearance of photospheric, chromospheric and coronal structures, including active regions and flares, seen in enhanced emissions in H alpha and different lines in the ultraviolet and extreme-ultraviolet as well as in white light observations, provides indications of the prevalent nature and importance of solar magnetic fields. Consequently, to understand the physics of active regions, the storage and release of flare energy, and the formation of hot plasmas and mass ejections, it is imperative that we understand and study the evolution of the Sun's magnetic field. To achieve such a goal, the logical first step is to seek a realistic representation of the configuration of the solar magnetic field. Current efforts in modeling physical structures and the storage and release of energy in flares are based on linear - so-called "constant-alpha" - force-free models of magnetic fields (e.g., Nakagawa et al., 1971; Nakagawa and Raadu, 1972; Welleck and Nakagawa, 1973). For example, Tanaka and Nakagawa (1973) used a linear force-free model to analyze the energy buildup for the August 1972 flare.

More recently, Schmahl et al. (1982) used a linear force-free model together with solar magnetograph data, VLA microwave maps, and X-ray spectroheliograms to study the evolution of an active region's magnetic structure, a study that led to a better understanding of the observed microwave structures. The authors concluded that localized currents must have

been present in the low corona to account for the bright 6 cm sources observed far from areas of strong sunspot fields, thus suggesting the presence of nonlinear (non-constant α) force-free fields. Further evidence for the existence of nonlinear force-free fields comes from the study of Krall et al. (1982); they used a linear model to investigate the vector magnetic field evolution within a flare-productive active region and concluded that the constant- α force-free model could not adequately represent the structures observed in the magnetic field of the active region. Previously, Levine (1976) had shown that changes in the value and sign of α can occur within a single active region.

The inadequacy of linear force-free models to represent observed solar magnetic fields is demonstrated amply by these and other studies. A compromise approach - constructing a "patchwork quilt" representation of the field of an active region by combining fields derived from solutions of the linear equations for different values of α - is of questionable value because it has no mathematical basis. It is certainly inappropriate in describing the evolution of magnetic fields when important nonlinear physical processes such as energy storage and release and MHD instabilities are involved.

In this paper a numerical solution is presented for extrapolating nonlinear force-free magnetic fields from a source surface, i.e., from observed vector magnetic fields at the photospheric level. The mathematical formalism of this method is presented in Section 2, and the numerical method and procedures are included in Section 3. In Section 4, results from a test

case are given together with an analysis of the computational accuracy for this test case. Finally, concluding remarks are presented in the last section.

2. Mathematical Model

The basic equation describing a force-free magnetic field is given by

$$\underline{j} \times \underline{B} = 0. \quad (1)$$

This may be rewritten with the aid of Ampere's Law (cgs electromagnetic units),

$$4 \pi \underline{j} = \nabla \times \underline{B}, \quad (2)$$

as

$$\nabla \times \underline{B} = \alpha \underline{B}, \quad (3)$$

where \underline{j} is the electric current density and alpha (α) in general is different for each field line, although it must be constant along a given field line.

This can be seen by taking the divergence of Equation (3) to obtain

$$\underline{B} \cdot \nabla \alpha = 0 \quad (4)$$

by virtue of the solenoidal condition

$$\nabla \cdot \underline{B} = 0. \quad (5)$$

If $\alpha = 0$, the field is potential; that is, the lowest-order approximation for a description of solar magnetic fields. Since a potential configuration represents the lowest state of energy of a given magnetic field, it is definitely not an appropriate description for magnetic fields in active regions that produce flares. If α has the same value throughout the field

domain, the resulting subclass of force-free fields is called a "constant α " or linear field, since the field components satisfy a linear differential equation (Nakagawa and Raadu, 1972).

In scalar form, Equations (3) and (5) form the basis of a scheme to extrapolate the force-free field when the vector magnetic field on the boundary surface is known. To show this, we write the following equations using Equations (3), (4) and (5):

$$\frac{\partial B_x}{\partial z} = \alpha B_y + \frac{\partial B_z}{\partial x}, \quad (6)$$

$$\frac{\partial B_y}{\partial z} = -\alpha B_x + \frac{\partial B_z}{\partial y}, \quad (7)$$

$$\frac{\partial B_z}{\partial z} = -\frac{\partial B_x}{\partial x} - \frac{\partial B_y}{\partial y}, \quad (8)$$

$$\alpha = \frac{1}{B_z} \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} \right), \quad (9)$$

$$\frac{\partial \alpha}{\partial z} = \frac{-1}{B_z} \left(B_x \frac{\partial \alpha}{\partial x} + B_y \frac{\partial \alpha}{\partial y} \right). \quad (10)$$

We take as the lower boundary surface the plane $z = 0$. Then Equation (9) indicates that knowing the vector field $\underline{B}(x,y,0)$ over this surface is

sufficient to determine the parameter $\alpha(x,y,0)$. Then, with $\alpha(x,y,0)$ and $B(x,y,0)$ specified, Equations (6) through (8) determine $\frac{\partial B}{\partial z}(x,y,0)$ and thus allow the start of an integration upward with height z . The process can then be iterated, beginning with the determination of $\alpha(x,y,dz)$ from $B(x,y,dz)$, again through Equation (9). It should be noted that Equation (10), derived from Equations (6) through (9), provides an alternate method to derive $\alpha(x,y,z>0)$. In places where B_z goes to zero, for example, along the "neutral line" in the photosphere (loci of nulls in the line-of-sight - B_z - component of the photospheric field), or near the tops of magnetic loops higher in the atmosphere, Equations (9) and (10) cannot be used. In these instances, an averaging procedure such as given below in Equation (24) is used to determine the value of α in Equation (9); a similar procedure is used to calculate $\frac{\partial \alpha}{\partial z}$. This is based on the assumption that all the field lines are continuous in the neighborhood of a point, an assumption that has a physical and mathematical basis since the present formulation does not include dissipative processes.

3. Numerical Methods

As pointed out by Grad and Rubin (1958), the differential equation for the force-free field problem is a mixed type, having two distinct real characteristics as in a hyperbolic equation and two imaginary ones as in the case of an elliptic equation. For the general nonlinear case this leads to unique mathematical difficulties both in the specification of boundary conditions and in the nature of the solutions. A number of astrophysical

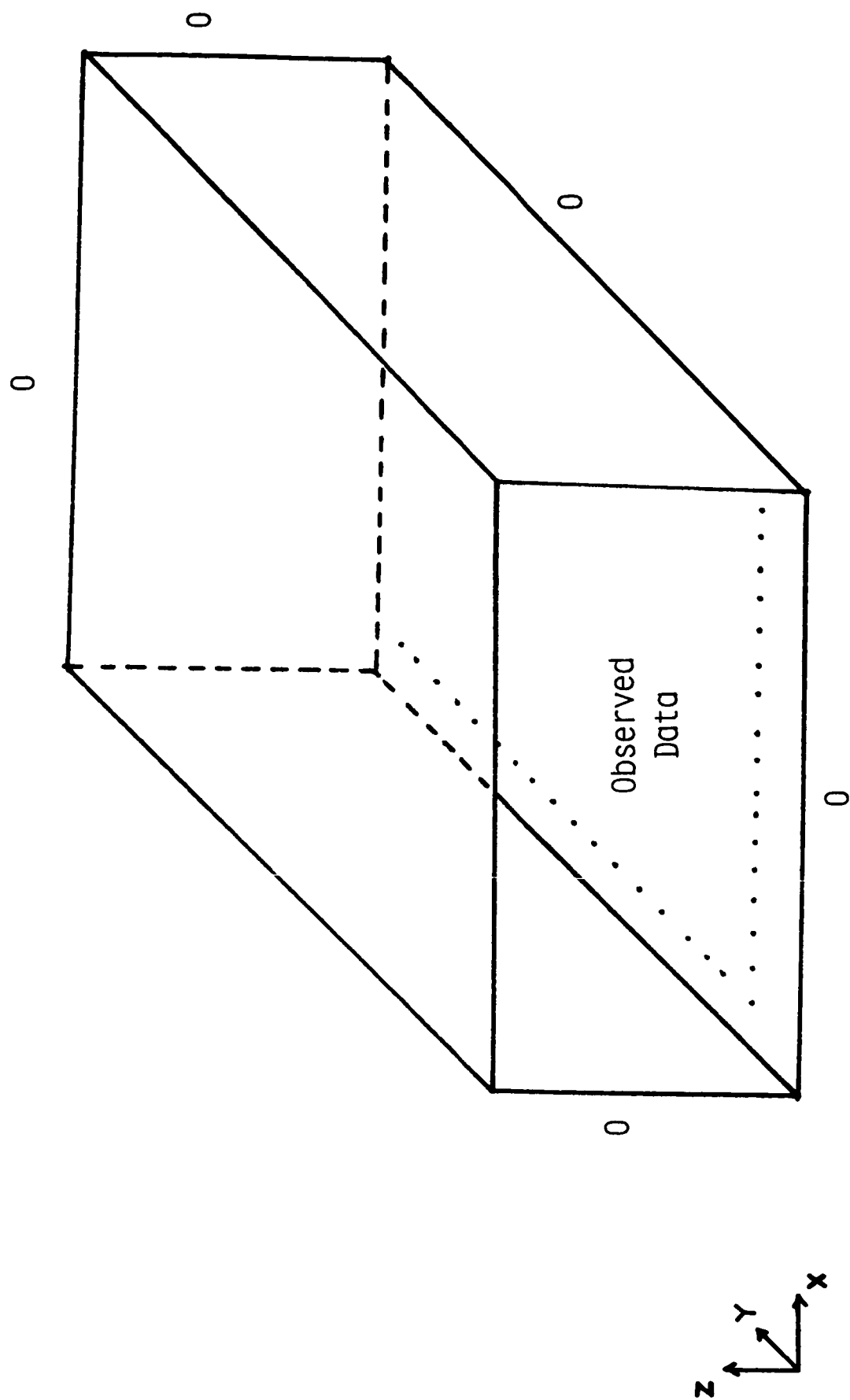


Fig. 1. The coordinate system and computational domain.

examples (magnetostatic as well as force-free) have been discussed by Low (1982a) and by Lerche and Low (1982). Recently several attempts have been made to devise algorithms for calculating nonlinear force-free fields using observational boundary conditions (e.g., Sakurai, 1981; Pridmore-Brown, 1981; Hannakam et al., 1984). The technique developed by Pridmore-Brown (1981) requires that the connectivity of the field lines be specified explicitly, although there is no established technique for determining this connectivity from the observational data. On the other hand, the method developed by Sakurai (1981) is a combination of the superposition of a current field on a potential field and a convergent iterative procedure. Neither Sakurai's or Pridmore-Brown's method has been used with observational data, nor have their methods been tested against nonlinear analytical models.

The numerical algorithm presented in this paper is a straightforward extrapolation procedure with specified boundary conditions. The computational procedures, differencing schemes and criteria for numerical stability are discussed in the following subsections.

3.1. COMPUTATIONAL PROCEDURES

We take as boundaries the six planes of the computational domain as shown in Figure 1. The values of the field on the lower surface are assumed known, e.g., from measurements of the vector magnetic field at the photospheric level. Values of the field on the other five surfaces are needed also to fully specify the boundary conditions.

The computational procedure to be used is summarized as follows:

1). At the lower surface ($z = 0$), the vector field (B_x, B_y, B_z) is prescribed at each grid point;

2). At this surface (i.e., $z = 0$), the horizontal derivatives

$$\frac{\partial B_x}{\partial x, y}, \quad \frac{\partial B_y}{\partial x, y}, \quad \frac{\partial B_z}{\partial x, y}$$

are computed;

3). Using Equation (9), the value of α over the plane $z = 0$ is computed;

4). Using the results from steps 2. and 3., the vertical derivatives are computed from Equations (6), (7) and (8).

5). With the vertical derivatives of $B(x, y, 0)$ thus determined, the field $B(x, y, dz)$ is computed using an explicit extrapolation scheme;

6). Repeating steps 2. through 5., the complete field configuration can be determined subject to the boundary conditions specified on the other five surfaces.

3.2. THE NUMERICAL DIFFERENTIAL SCHEME

In order to compute these horizontal derivatives numerically, we used second order central differences for the interior points (Burden, 1981), viz.

$$\left(\frac{\partial B_i}{\partial x} \right)_{l,m,n} = \frac{(B_i)_{l+1,m,n} - (B_i)_{l-1,m,n}}{2 h_x} + O(h_x^2), \quad (11)$$

and

$$\left(\frac{\partial B_i}{\partial y} \right)_{l,m,n} = \frac{(B_i)_{l,m+1,n} - (B_i)_{l,m-1,n}}{2 h_y} + O(h_y^2), \quad (12)$$

where the subscript "i" represents the components of the physical quantities (i.e., $i = x, y, z$), l, m , and n indicate the coordinates of the grid points in a three-dimensional space, and h_x and h_y are the grid spacings along the x and y axes. These grid spacings are chosen according to the numerical stability criteria which are discussed in the next subsection.

For the computation of derivatives for points on the boundaries, the central differences cannot be used; the following expressions were used in these cases:

for points $l = 1$,

$$\left(\frac{\partial B_i}{\partial x} \right)_{l,m,n} = \frac{-(B_i)_{l+2,m,n} + 4(B_i)_{l+1,m,n} - 3(B_i)_{l,m,n}}{2 h_x} + O(h_x^2), \quad (13)$$

for points $m = 1$,

$$\left(\frac{\partial B_i}{\partial y} \right)_{l,m,n} = \frac{-(B_i)_{l,m+2,n} + 4(B_i)_{l,m+1,n} - 3(B_i)_{l,m,n}}{2 h_y} + O(h_y^2); \quad (14)$$

for points $l = l_{\max}$,

$$\left(\frac{\partial B_i}{\partial x} \right)_{l,m,n} = \frac{3(B_i)_{l,m,n} - 4(B_i)_{l-1,m,n} + (B_i)_{l-2,m,n}}{2 h_x} + O(h_x^2), \quad (15)$$

for points $m = m_{\max}$

$$\left(\frac{\partial B_i}{\partial y} \right)_{l,m,n} = \frac{3(B_i)_{l,m,n} - 4(B_i)_{l,m-1,n} + (B_i)_{l,m-2,n}}{2 h_y} + O(h_y^2). \quad (16)$$

Finally, to numerically extrapolate the field components in the z direction using the derived vertical gradients, both Euler's formula and the Adams-Bashforth two-step formula were used. These formulas are:

Euler's formula,

$$(B_i)_{l,m,n+1} - (B_i)_{l,m,n} = h_z \left(\frac{\partial B_i}{\partial z} \right)_{l,m,n} + O(h_z^2) ; \quad (17)$$

Adams-Bashforth,

$$(B_i)_{l,m,n+1} - (B_i)_{l,m,n} = \frac{3h_z}{2} \left(\frac{\partial B_i}{\partial z} \right)_{l,m,n} - \frac{h_z}{2} \left(\frac{\partial B_i}{\partial z} \right)_{l,m,n-1} + O(h_z^2) . \quad (18)$$

Thus, the final numerical forms for the extrapolation of the magnetic field for points interior to the domain are given by combining Equations (6)-(9) and (11),(12) in Equation (17):

$$(B_x)_{l,m,n+1} = (\bar{B}_x)_{l,m,n} + \frac{h_z}{2h_x} \left[(B_z)_{l+1,m,n} - (B_z)_{l-1,m,n} \right] + (\alpha B_y)_{l,m,n} h_z , \quad (19)$$

$$(B_y)_{l,m,n+1} = (\bar{B}_y)_{l,m,n} + \frac{h_z}{2h_y} [(B_z)_{l,m+1,n} - (B_z)_{l,m-1,n}] - (\alpha B_x)_{l,m,n} h_z, \quad (20)$$

$$(B_z)_{l,m,n+1} = (B_z)_{l,m,n} - \frac{h_z}{2h_x} [(B_x)_{l+1,m,n} - (B_x)_{l-1,m,n}] - \frac{h_z}{2h_y} \bullet \quad (21)$$

$$[(B_y)_{l,m+1,n} - (B_y)_{l,m-1,n}],$$

$$(\alpha)_{l,m,n+1} = (\bar{\alpha})_{l,m,n+1} + \frac{1}{(B_z)_{l,m,n+1}} \left[\frac{(B_y)_{l+1,m,n+1} - (B_y)_{l-1,m,n+1}}{2 h_x} - \frac{(B_x)_{l,m+1,n+1} - (B_x)_{l,m-1,n+1}}{2 h_y} \right], \quad (22)$$

where

$$(\bar{B}_i)_{l,m,n} = \frac{1}{4} [(B_i)_{l+1,m,n} + (B_i)_{l-1,m,n} + (B_i)_{l,m+1,n} + (B_i)_{l,m-1,n}], \quad (23)$$

and

$$(\bar{\alpha})_{l,m,n} = \frac{1}{4} [(\alpha)_{l+1,m,n} + (\alpha)_{l-1,m,n} + (\alpha)_{l,m+1,n} + (\alpha)_{l,m-1,n}] \quad (24)$$

These averaging formulas for $(\alpha)_{l,m,n}$ and $(B_i)_{l,m,n}$ are used instead of the values of α and B_i at the grid points as a method for smoothing the data.

3.3. BOUNDARY CONDITIONS AND NUMERICAL STABILITY

In general, the boundary conditions imposed at the six planes forming the boundary of the computational domain are specified. The values on the lower surface are determined from observational data; thus this surface is referred to as the "source surface." Values on the other surfaces are prescribed according to both physical and mathematical conditions. In this paper, we employ the following representative boundary conditions:

- (i). As $z \rightarrow \infty$, $(B_x, B_y, B_z) \rightarrow 0$.
- (ii). Outside the rectangle given by $-a < x < a$, $-b < y < b$, periodic conditions are chosen.

In order to assure numerical stability, we have used von Neumann's method (Mitchell and Griffiths, 1980), which leads to the following conditions:

$$\frac{B_x}{B_z} \delta_x \leq \frac{1}{2} \quad \text{and} \quad \frac{B_y}{B_z} \delta_y \leq \frac{1}{2}, \quad (25)$$

$$\frac{B_y^2 + B_z^2}{B_x B_z} \delta_x \leq \frac{1}{2} \quad \text{and} \quad \frac{B_x^2 + B_z^2}{B_y B_z} \delta_y \leq \frac{1}{2}, \quad (26)$$

where

$$\delta_x = \frac{h_z}{h_x} \quad , \quad \delta_y = \frac{h_z}{h_y} \quad (27)$$

The detailed derivation of these criteria is given in the Appendix. These stability criteria guide us in choosing the proper grid spacings and extrapolation step (dz).

4. Computational Results

In order to demonstrate this algorithm, we have carried out a numerical calculation using as a test case a form for the nonlinear force-free field that admits of an analytical solution. The form chosen was the analytical solution of Low (1982b) that incorporates a distorted magnetic neutral line and a highly sheared transverse magnetic field in the vicinity of the neutral line, both important elements in flare-productive active regions (Hagyard et al., 1984; Hagyard, Moore and Emslie, 1984). Low's solution is given by the following equations:

$$B_x = - B_0 \frac{\cos \phi_i}{r} \quad , \quad (28)$$

$$B_y = \frac{B_0 (x+x_0) (y+y_0)}{r R_i} \cos \phi_i - \frac{B_0 (z+z_0)}{R_i} \sin \phi_i \quad , \quad (29)$$

$$B_z = \frac{B_0 (x+x_0) (z+z_0)}{r R_i} \cos \phi_i + \frac{B_0 (y+y_0)}{R_i} \sin \phi_i , \quad (30)$$

where

$$R_i^2 = [(y + y_0)^2 + (z + z_0)^2] , \quad (31)$$

$$r^2 = [(x + x_0)^2 + (y + y_0)^2 + (z + z_0)^2] , \quad (32)$$

$$\phi_i = 0.3 * \ln(R) , \quad (33)$$

$$\alpha = \frac{d \phi_i(r)}{dr} = - 0.3/R , \quad (34)$$

B_0 is the magnitude of the magnetic field strength at the origin (x_0, y_0, z_0) , and R is the normalized variable r ($R = r/R_0$).

Using these formulas, we generated the values of the magnetic field on the source surface and the other five surfaces forming the boundary of the domain of the calculation. These numerical values were then used in our numerical algorithm to extrapolate the field above the source surface. Since the analytic solutions gave the exact solutions for the field above the source surface, comparisons of our results with the analytic computations provided a stringent test of the algorithm. Figures 2-5 show the resulting comparisons between the numerical and analytic solutions.

B_z CONTOURS ON X-Y PLANE AT A SPECIFIC HEIGHT ($Z = \text{CONST.}$)

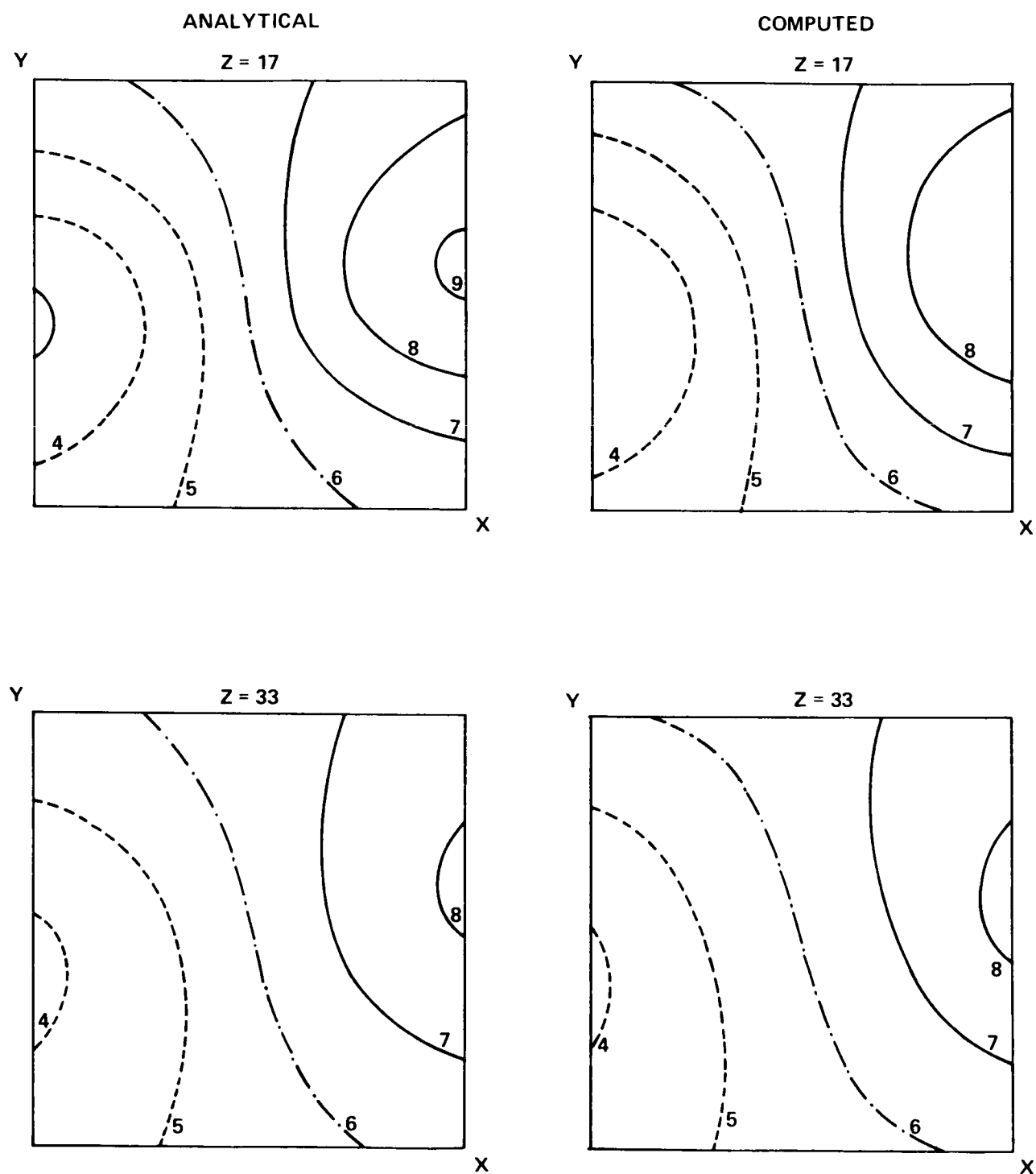


Fig. 2. Comparison of the analytical and numerical results for the vertical component of the magnetic field at two different heights z . The solid and dashed curves represent positive and negative contours of the vertical (B_z) component of the nonlinear force-free magnetic field in the x-y plane at $z = 17$ (in the middle of the z scale) and at $z = 33$ (near the top of the z scale).

VECTORAL REPRESENTATION OF MAGNETIC FIELD LINES PROJECTED ON X-Y PLANE
AT A SPECIFIC HEIGHT ($Z = \text{CONST.}$)

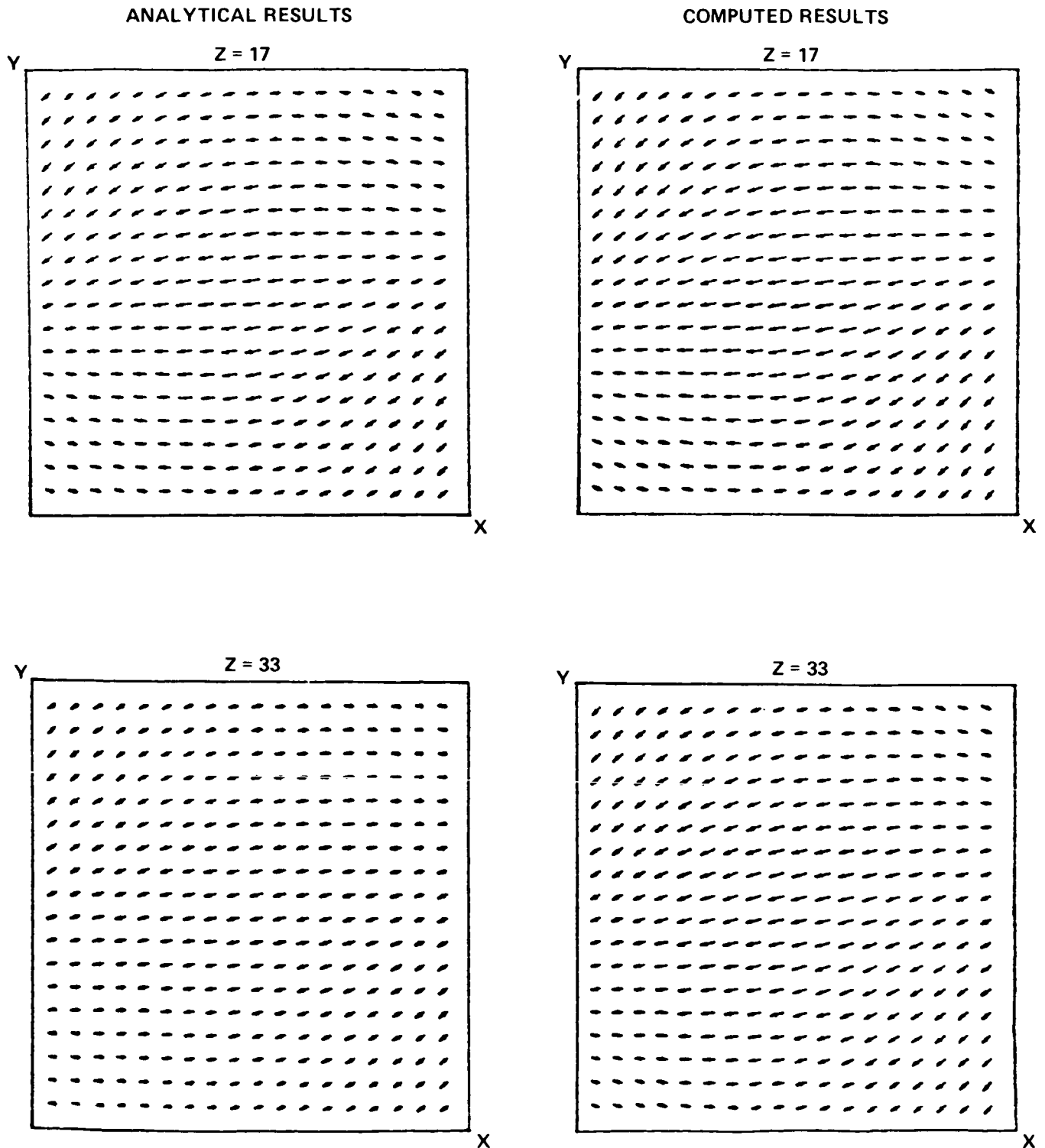


Fig. 3. Comparison of the analytical and numerical results for the transverse component of the magnetic field at two different heights z . The magnitude and direction of the vector field projected in the x-y (transverse) plane are represented by the length and direction of the line segments at each grid point.

VECTORAL REPRESENTATION OF MAGNETIC FIELD LINES PROJECTED ON
X-Z PLANE AT A SPECIFIC Y POSITION

ANALYTICAL RESULTS

COMPUTED RESULTS

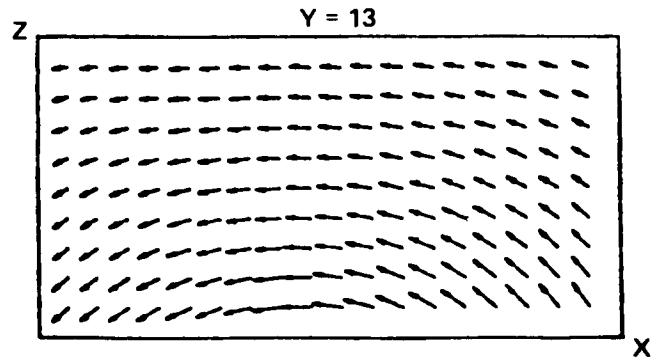
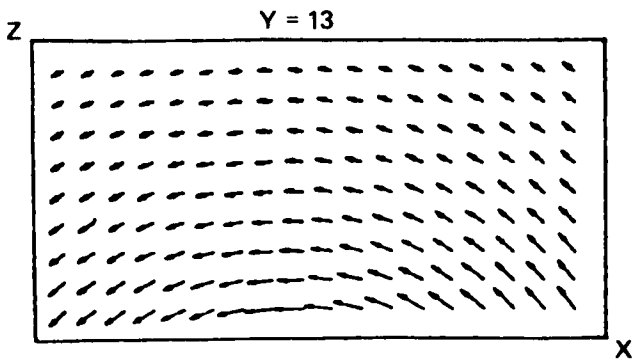
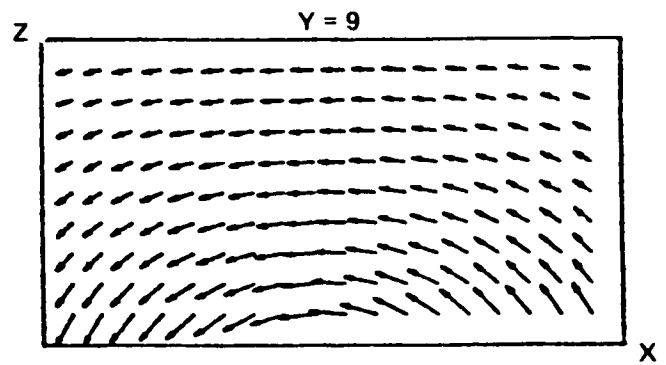
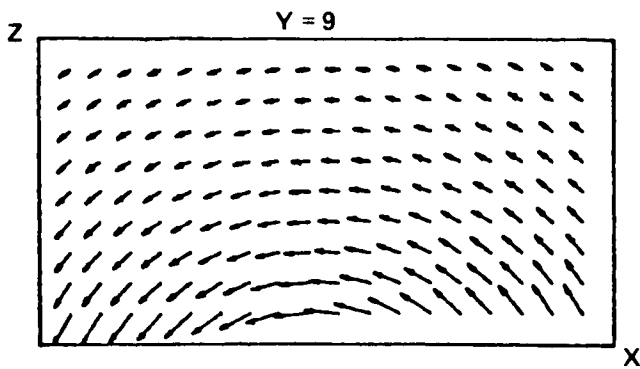
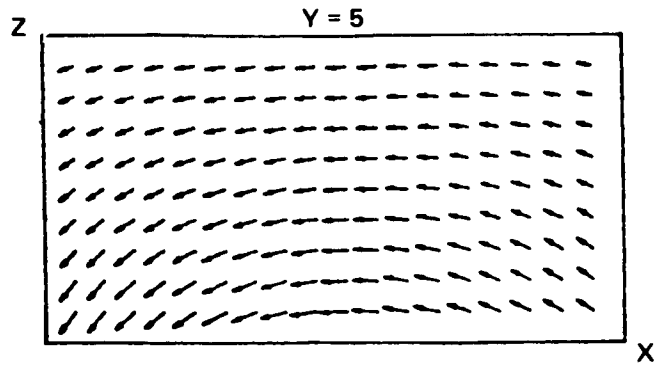
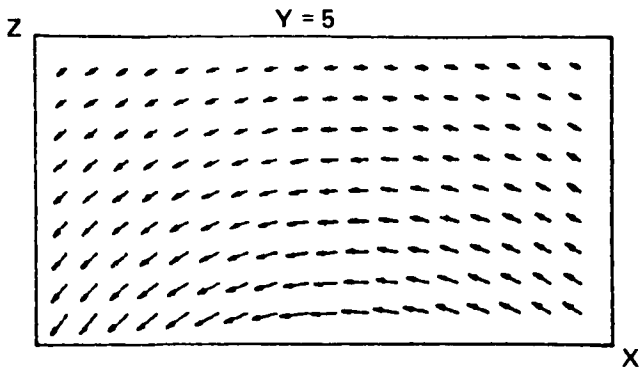


Fig. 4. Comparison of the analytical and numerical results for the vector magnetic field projected on the x-z (vertical) plane. The locations of the planes $y = 5, 9$ and 13 can be seen in Figure 7.

VECTORAL REPRESENTATION OF MAGNETIC FIELD LINES PROJECTED ON
Y-Z PLANE AT A SPECIFIC X POSITION

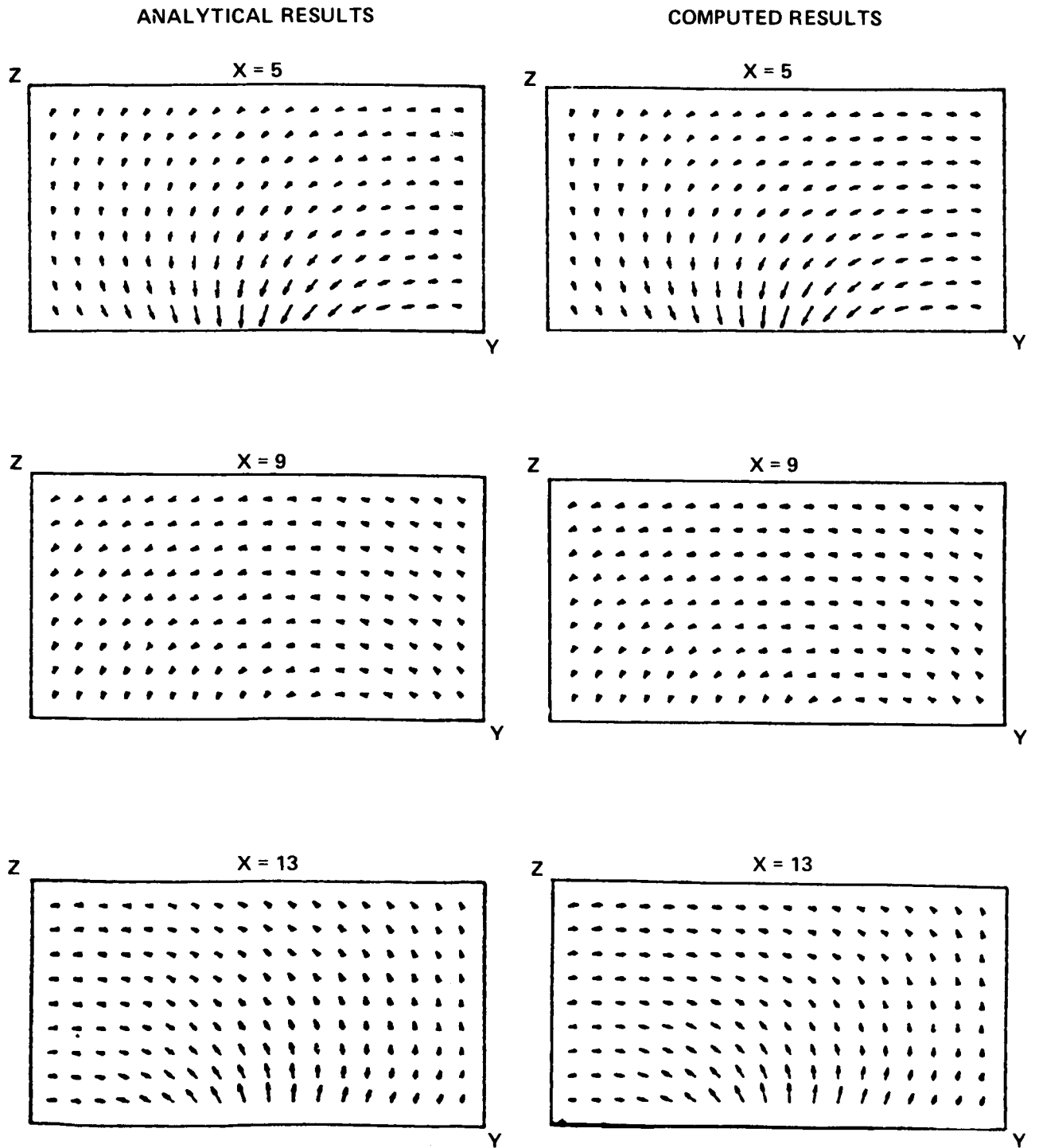


Fig. 5. Comparison of the analytical and numerical results for the vector magnetic field projected on the y-z (vertical) plane. The locations of the planes $x = 5, 9$ and 13 can be seen in Figure 7.

In Figure 2, contours of the vertical field (the line-of-sight field for areas near the center of the solar disk) are shown in the x-y plane at two heights ($z = \text{constant}$) above the photosphere. Normalized to the computational size of the square base region at $z = 0$, the two levels are at $z = 0.50$ and 1.0 ; with a base length of $\sim 10^5$ km, these levels correspond to heights of 50,000 and 100,000 km, respectively. Since these represent considerable heights up into the solar corona, comparisons between the numerical and analytical fields at these heights represent critical tests of the numerical method. As judged by the results shown in Figure 2, there is good agreement at both levels as far as the strength of the vertical component of the magnetic field is concerned.

It is also important to examine the strength and direction of the transverse components of the field vectors at specified locations. In Figure 3 we show a comparison between the analytic and computed transverse field in the x-y plane at two different levels, again $z = 17$ and $z = 33$. Figures 4 and 5 make the same comparisons for two more perspective views of the field vector: in the two vertical planes x-z and y-z. In all three cases the qualitative agreement is good.

In Table 1, we present quantitative comparisons for the height $z = 33$. For alternating grid points in the x-y plane, the table gives the analytic and computed values for all three field components and for the force-free parameter α . From data such as these in Table 1, we have calculated the mean-square percentage errors in the field components at different heights as

TABLE I.

TABULATED VALUES OF B_X , B_Y , B_Z AND α AT $Z = 33$ FOR ANALYTICAL AND COMPUTED RESULTS

N3 = 33		N2 =		7		11		15		19		23		27		31	
	N1	ANAL- YTICAL		COM- PUTED		ANAL- YTICAL		COM- PUTED		ANAL- YTICAL		COM- PUTED		ANAL- YTICAL		COM- PUTED	
		7		11		15		19		23		27		31			
7	BX	-.238	-.250	-.250	-.253	-.258	-.256	-.262	-.257	-.260	-.254	-.254	-.250	-.244	-.250	-.250	
	BY	-.056	-.083	-.072	-.092	-.091	-.098	-.111	-.098	-.128	-.100	-.140	-.107	-.145	-.108	-.108	
	BZ	-.141	-.151	-.143	-.152	-.140	-.151	-.129	-.143	-.110	-.123	-.085	-.106	-.058	-.103	-.103	
	AL	-.078	-.088	-.081	-.076	-.083	-.063	-.085	-.048	-.084	-.032	-.082	-.020	-.079	-.010	-.010	
11	BX	-.250	-.238	-.263	-.246	-.273	-.250	-.277	-.251	-.276	-.251	-.268	-.247	-.257	-.239	-.239	
	BY	-.065	-.078	-.078	-.086	-.093	-.094	-.108	-.097	-.120	-.101	-.128	-.111	-.130	-.113	-.113	
	BZ	-.114	-.130	-.113	-.121	-.107	-.112	-.096	-.099	-.078	-.080	-.056	-.066	-.032	-.072	-.072	
	AL	-.081	-.092	-.085	-.084	-.088	-.074	-.089	-.061	-.089	-.045	-.086	-.031	-.083	-.020	-.020	
15	BX	-.258	-.226	-.273	-.242	-.283	-.253	-.288	-.257	-.286	-.256	-.279	-.248	-.266	-.231	-.231	
	BY	-.077	-.079	-.087	-.087	-.097	-.094	-.106	-.098	-.113	-.102	-.116	-.110	-.116	-.109	-.109	
	BZ	-.084	-.094	-.079	-.080	-.070	-.066	-.057	-.051	-.040	-.035	-.021	-.026	-.002	-.031	-.031	
	AL	-.083	-.095	-.088	-.090	-.091	-.082	-.092	-.070	-.092	-.054	-.089	-.043	-.086	-.028	-.028	
19	BX	-.262	-.219	-.277	-.241	-.288	-.257	-.293	-.264	-.291	-.262	-.283	-.250	-.270	-.227	-.227	
	BY	-.091	-.087	-.097	-.096	-.103	-.099	-.106	-.100	-.107	-.101	-.105	-.103	-.101	-.100	-.100	
	BZ	-.053	-.048	-.042	-.038	-.030	-.027	-.015	-.013	-.001	-.003	-.017	-.014	-.031	-.018	-.018	
	AL	-.085	-.084	-.089	-.097	-.092	-.083	-.094	-.089	-.093	-.077	-.091	-.078	-.087	-.079	-.079	
23	BX	-.260	-.219	-.276	-.241	-.286	-.256	-.291	-.261	-.290	-.259	-.282	-.249	-.268	-.231	-.231	
	BY	-.106	-.095	-.110	-.106	-.111	-.105	-.108	-.100	-.103	-.097	-.096	-.094	-.088	-.088	-.088	
	BZ	-.022	-.003	-.006	-.006	-.011	-.010	-.028	-.023	-.043	-.039	-.056	-.053	-.065	-.066	-.066	
	AL	-.084	-.048	-.089	-.059	-.094	-.067	-.093	-.074	-.093	-.076	-.090	-.081	-.086	-.095	-.095	
27	BX	-.254	-.228	-.268	-.243	-.279	-.251	-.283	-.253	-.282	-.252	-.274	-.247	-.262	-.238	-.238	
	BY	-.122	-.100	-.123	-.111	-.120	-.106	-.112	-.099	-.101	-.096	-.088	-.090	-.076	-.082	-.082	
	BZ	-.007	-.048	-.028	-.047	-.050	-.030	-.069	-.066	-.084	-.082	-.093	-.095	-.098	-.108	-.108	
	AL	-.082	-.022	-.086	-.035	-.089	-.040	-.091	-.059	-.090	-.073	-.088	-.083	-.085	-.091	-.091	
31	BX	-.244	-.239	-.257	-.246	-.266	-.250	-.270	-.253	-.268	-.253	-.262	-.251	-.251	-.247	-.247	
	BY	-.137	-.102	-.137	-.110	-.130	-.103	-.117	-.098	-.101	-.097	-.083	-.093	-.067	-.085	-.085	
	BZ	-.033	-.079	-.059	-.082	-.084	-.092	-.105	-.113	-.120	-.129	-.127	-.136	-.127	-.140	-.140	
	AL	-.079	-.011	-.083	-.021	-.086	-.032	-.087	-.048	-.086	-.062	-.085	-.075	-.081	-.085	-.085	

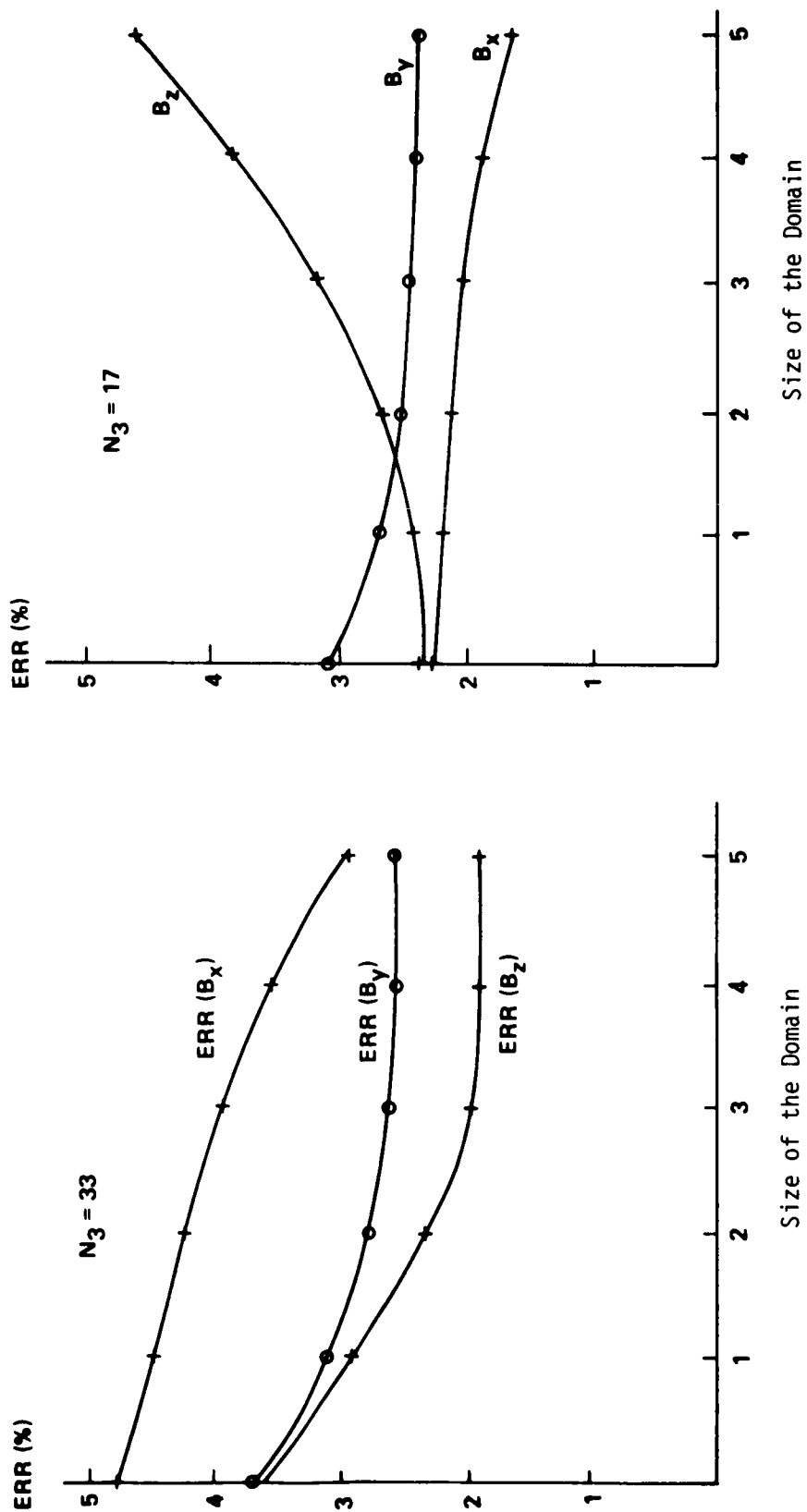


Fig. 6. The average percentage error between the analytical and numerical solutions as a function of the parameter δ . This parameter represents how much of the domain in the x-y plane is included in the analysis of average errors. $\delta = 0$ represents the initial domain; $\delta = 1$ indicates that the first row and column at the boundary of the domain are omitted in the error analysis, etc.

a function of the parameter δ which is the number of rows and columns next to the boundary surfaces in the x-y plane that are omitted from the error analysis. In Figure 6, these percentage errors are plotted against the parameter δ for the levels $z = 17$ and $z = 33$. For the case of $\delta = 0$, that is, if all grid points in the domain are included in the error analysis, the basic quantitative result is very obvious: at all height levels, including the highest, the computed solution agrees with the analytic solution to better than 5% in all field components, and the typical error is only 3%. Figure 6 also demonstrates that the worst errors tend to occur near the boundaries since the errors at $z = 33$ all tend to decrease as the grid points near the boundary are omitted in the error calculation. This result indicates that in computations using observational data some care must be exercised in analyzing the data near the boundaries.

5. Concluding Remarks

In this paper, a numerical scheme was presented for the extrapolation of nonlinear force-free magnetic fields from a given source surface. This newly developed algorithm was tested by using an analytical solution to the nonlinear force-free field equations; the test showed that, with the present numerical scheme, nonlinear force-free magnetic fields can be extrapolated to an accuracy of better than 5% in comparison with the analytical case. As has been pointed out by a number of authors (Grad, 1973; Kress, 1977,1978), the complete mathematical characterization of the nonlinear force-free problem has not yet been achieved. Accordingly, we do not claim that we have solved the

nonlinear force-free problem; we merely have presented a numerical algorithm that can be used for extrapolation of the force-free solution to within a certain accuracy. Nevertheless, because of the practical importance of force-free fields in understanding the physics of the Sun, there is strong motivation for devising a numerical approach that will serve in the interim until a satisfactory mathematical understanding is at hand. As of now, we have achieved this initial goal; the next step is to apply this algorithm to actual observations of vector magnetic fields to test the numerical code under less than ideal conditions, i.e., using noisy data. With data from the MSFC vector magnetograph (Hagyard et al., 1982), we expect to accomplish this next goal in the near future.

Acknowledgements

The work done by STW and HMC was partially supported by a NASA grant (NAGW - 9) and a NASA/MSFC contract (NAS8-33525). STW was also partially supported by an AFGL grant (F19628-83-k-0019).

References

- Burden, R.L.: 1981, Numerical Analysis, Prindle, Weber and Schmidt Publ. Co., Boston.
- Grad, H.: 1973; Adv. Plasma Phys. 5, 103.
- Grad, H., and Rubin, H.: 1958, in Proc. Second Intn'l. Conf. on the Peaceful Uses of Atomic Energy, Geneva.
- Hagyard, M.J., Cumings, N.P., West, E.A., and Smith, J.E.: 1982, Solar Phys. 80, 33.
- Hagyard, M.J., Smith, J.B.Jr., Teuber, D., and West, E.A.: 1984, Solar Phys. 91, 115.
- Hagyard, M.J., Moore, R.L., and Emslie, A.G.: 1984, Advances in Space Research, in press.
- Hannakam, L., Gary, G. Allen, and Teuber, D.L.: 1984, Solar Phys., in press.
- Krall, K.R., Smith, J.B. Jr., Hagyard, M.J., West, E.A., and Cumings, N.P.: 1982, Solar Phys. 79, 59.
- Kress, R.: 1977, Z. Angew. Math. Phys. 28, 715.
- Kress, R.: 1978, Proc. Roy. Soc. Edinburgh 82A, 71.
- Lerch, I. and Low, B.C.: 1982, Physica 4D, 293.
- Levine, R.: 1976, Solar Phys. 46, 159.
- Low, B.C.: 1982a, Rev. Geophy. Space Phys. 20, 145.
- Low, B.C.: 1982b, Solar Phys. 77, 43.
- Mitchell, A.R., and Griffiths, D.F.: 1980, The Finite Difference Method in Partial Differential Equations, Wiley, New York.

- Nakagawa, Y., Raadu, M.A., Billings, D.F., and McNamara, D.: 1971, Solar Phys. 19, 72.
- Nakagawa, Y., and Raadu, M.A.: 1972, Solar Phys. 25, 127.
- Pridmore-Brown, D.C.: 1981, Aerospace Corporation Report No. ATR-81(7813)-1.
- Sakurai, T.: 1981, Solar Phys. 69, 343.
- Schmahl, E.J., Kundu, M.R., Strong, K.T., Bentley, R.D., Smith, J.B. Jr., and Krall, K.R.: 1982, Solar Phys. 80, 233.
- Sturrock, P.A., and Woodbury, E.T.: 1967, in Plasma Astrophysics, ed. P.A. Sturrock, Academic Press, New York, 155.
- Tanaka, K., and Nakagawa, Y.: 1973, Solar Phys. 33, 187.
- Welleck, R.E., and Nakagawa, Y.: 1973, NCAR - TN/STR-87.

Appendix: Derivation of the Numerical Stability Criteria for the Vertical Direction

According to our extrapolation scheme, the i^{th} component of the magnetic field at a particular point in space ($x=l$, $y=m$, $z=n+1$) can be written to the first order of approximation in the following form:

$$(B_i)_{l,m,n+1} = (\bar{B}_i)_{l,m,n} + h_z \left(\frac{\partial B_i}{\partial z} \right)_{l,m,n}, \quad i=1,2,3,4, \quad (\text{A-1})$$

where

$$(\bar{B}_i)_{l,m,n} = \frac{1}{4} \left[(B_i)_{l-1,m,n} + (B_i)_{l+1,m,n} + (B_i)_{l,m-1,n} + (B_i)_{l,m+1,n} \right], \quad (\text{A-2})$$

and

$$B_1 = B_x, B_2 = B_y, B_3 = B_z \text{ and } B_4 = \alpha.$$

From the original set of governing equations (i.e., Equations (6), (7), (8), (9), and (10), we obtain

$$\frac{\partial B_1}{\partial z} = \frac{\partial B_3}{\partial x} + \frac{B_2}{B_3} \left(\frac{\partial B_2}{\partial x} - \frac{\partial B_1}{\partial y} \right), \quad (\text{A-3})$$

$$\frac{\partial B_2}{\partial z} = \frac{\partial B_3}{\partial y} - \frac{B_1}{B_3} \left(\frac{\partial B_2}{\partial x} - \frac{\partial B_1}{\partial y} \right), \quad (\text{A-4})$$

$$\frac{\partial B_3}{\partial z} = -\frac{\partial B_1}{\partial x} - \frac{\partial B_2}{\partial y} \quad , \quad (A-5)$$

$$\frac{\partial B_4}{\partial z} = -\frac{B_1}{B_3} \frac{\partial B_4}{\partial x} - \frac{B_2}{B_3} \frac{\partial B_4}{\partial y} \quad (A-6)$$

Using the Von Neuman method to derive the numerical stability criteria, we assume the i^{th} component of a physical quantity can be represented by

$$B_i = A_i T(z) \exp [i (\gamma_{ix} x + \gamma_{iy} y)] \quad , \quad (A-7)$$

where A_i , a positive constant, is the amplitude of B_i , the parameters γ_{ix} and γ_{iy} are positive phase constants, $i=(-1)^{\frac{1}{2}}$, and $T(z)$ is an arbitrary function of z . Therefore, we can represent Equation (A-7) at each grid point as

$$(B_i)_{l,m,n} = A_i T_n \exp [i (\gamma_{ix} l h_x + \gamma_{iy} m h_y)] \quad , \quad (A-8)$$

where h_x and h_y are the grid spacings in the x and y directions. Substituting Equation (A-8) into (A-2), we have

$$\frac{(\bar{B}_i)_{l,m,n}}{(B_i)_{l,m,n}} = \frac{1}{2} (\cos \beta_{ix} + \cos \beta_{iy}) \quad , \quad (A-9)$$

where

$$\beta_{ix} = \gamma_{ix} h_x \geq 0 \quad (A-10)$$

and

$$\beta_{iy} = \gamma_{iy} h_y \geq 0 \quad .$$

Substituting Equations (A-8), (A-9), and (A-2) into (A-1) yields

$$\frac{(B_i)_{l,m,n+1}}{(B_i)_{l,m,n}} = \frac{T_{n+1}}{T_n} = \frac{1}{2} (\cos \beta_{ix} + \cos \beta_{iy}) + \frac{h_z}{(B_i)_{l,m,n}} \left(\frac{\partial B_i}{\partial z} \right)_{l,m,n} \quad (A-11)$$

Employing central differences for the derivatives and using Equation (A-7), we can express the horizontal gradients in the form

$$\begin{aligned} \frac{1}{(B_i)_{l,m,n}} \left(\frac{\partial B_i}{\partial x} \right)_{l,m,n} &= \frac{1}{2 h_x} \left[\exp (-i \gamma_{ix} h_x) \right. \\ &\quad \left. - \exp (i \gamma_{ix} h_x) \right] = \frac{i}{h_x} \sin \beta_{ix} , \end{aligned} \quad (A-12)$$

and

$$\begin{aligned} \frac{1}{(B_i)_{l,m,n}} \left(\frac{\partial B_i}{\partial y} \right)_{l,m,n} &= \frac{1}{2 h_y} \left[\exp (-i \gamma_{iy} h_y) \right. \\ &\quad \left. - \exp (i \gamma_{iy} h_y) \right] = \frac{i}{h_y} \sin \beta_{iy} . \end{aligned} \quad (A-13)$$

Substituting Equations (A-12) and (A-13) into Equations (A-3) through (A-6), we obtain

$$\left(\frac{1}{B_1} \frac{\partial B_1}{\partial z} \right)_{1,m,n} = \frac{i B_3}{h_x B_1} \sin \beta_{3x} + \frac{i B_2^2}{h_x B_1 B_3} \sin \beta_{2x} - \frac{i B_2}{h_y B_3} \sin \beta_{1y} , \quad (\text{A-14})$$

$$\left(\frac{1}{B_2} \frac{\partial B_2}{\partial z} \right)_{1,m,n} = \frac{i B_3}{h_y B_2} \sin \beta_{3y} + \frac{i B_1}{h_x B_3} \sin \beta_{2x} + \frac{i B_1^2}{h_y B_2 B_3} \sin \beta_{1y} , \quad (\text{A-15})$$

$$\left(\frac{1}{B_3} \frac{\partial B_3}{\partial z} \right)_{1,m,n} = -i \left(\frac{B_1}{h_x B_3} \sin \beta_{1x} + \frac{B_2}{h_y B_3} \sin \beta_{2y} \right) , \quad (\text{A-16})$$

$$\left(\frac{1}{B_4} \frac{\partial B_4}{\partial z} \right)_{1,m,n} = -i \left(\frac{B_1}{h_x B_3} \sin \beta_{4x} + \frac{B_2}{h_y B_3} \sin \beta_{4y} \right) \quad (\text{A-17})$$

For simplicity, we assume

$$\beta_{ix} = \beta_x , \quad \beta_{iy} = \beta_y \quad \text{for } i = 1, 2, 3, 4,$$

and define

$$\frac{A_1}{A_3} = \frac{(B_1)_{1,m,n}}{(B_3)_{1,m,n}} , \quad (\text{A-18})$$

$$\frac{A_2}{A_3} = \frac{(B_2)_{1,m,n}}{(B_3)_{1,m,n}} , \quad (\text{A-19})$$

and

$$\delta_x = h_z/h_x , \quad \delta_y = h_z/h_y \quad (\text{A-20})$$

Using the results given in Equations (A-14) through (A-17), Equation (A-11) becomes

$$\frac{T_{n+1}}{T_n} = \frac{1}{2} (\cos \beta_x + \cos \beta_y) - i (L_{ix} \sin \beta_x + L_{iy} \sin \beta_y) , \quad (\text{A-21})$$

with $i=1,2,3,4$, and

$$L_{1x} = - \left(\frac{A_3^2 + A_2^2}{A_1 A_3} \right) \delta_x , \quad (A-22)$$

$$L_{2x} = L_{3x} = L_{4x} = \frac{A_1}{A_3} \delta_x , \quad (A-23)$$

$$L_{1y} = L_{3y} = L_{4y} = \frac{A_2}{A_3} \delta_y , \quad (A-24)$$

$$L_{2y} = - \left(\frac{A_3^2 + A_1^2}{A_2 A_3} \right) \delta_y \quad (A-25)$$

In order to have this numerical scheme be stable along the z-axis, we require the square of the function $\frac{T_{n+1}}{T_n}$ expressed in Equation (A-21) to be less than unity (Mitchell and Griffiths, 1980), i.e.,

$$\left| \frac{T_{n+1}}{T_n} \right|^2 = (L_{ix} \sin \beta_x + L_{iy} \sin \beta_y)^2 - \frac{1}{4} (\sin \beta_x + \sin \beta_y)^2 - \left[\sin \left(\frac{\beta_x - \beta_y}{2} \right) \right]^2 + 1 < 1 , i = 1,2,3,4 . \quad (A-26)$$

This implies the following inequality must hold:

$$(L_{ix} \sin \beta_x + L_{iy} \sin \beta_y)^2 \leq \frac{1}{4} (\sin \beta_x + \sin \beta_y)^2 + \left[\sin \left(\frac{\beta_x - \beta_y}{2} \right) \right]^2 \quad (A-27)$$

(A-27)

One can show that this relation is satisfied for L_{ix}^2 , L_{iy}^2 , L_{ix} .

$L_{iy} \leq \frac{1}{4}$, or $|L_{ix}|$, $|L_{iy}| \leq 1/2$, which leads to the stability criteria in Equations (27) and (28).

Modeling Solar Magnetic Structures

B. C. Low

High Altitude Observatory
National Center for Atmospheric Research¹

Abstract. I will discuss some ideas in the theoretical study of force-free magnetic fields and magnetostatic fields, which are relevant to the effort of using magnetograph data as inputs to model the quasi-static, large-scale magnetic structures in the solar atmosphere. Basic physical principles will be emphasized. An attempt will be made to assess what we may learn, physically, from the models based on these ideas. There is prospect for learning useful physics and this ought to be an incentive for intensifying the efforts to improve vector magnetograph technology and to solve the basic radiative-transfer problems encountered in the interpretation of magnetograph raw data.

1. Introduction

It is now nearly twenty years that we have been building large-scale magnetic-field models based on magnetograph measurements. I will review some theoretical ideas of relevance to this kind of work. The review covers selected topics in the theoretical study of force-free magnetic fields and the more general magnetostatic fields, which, in my opinion, are likely to prove useful. I will emphasize the basic principles and attempt to assess the usefulness of the models that can be constructed, in terms of how much physics we can learn from them.

The quantitative measurement of solar magnetic fields began with the instrument built by Babcock and Babcock (1955) which registered the first order, circularly polarized, Zeeman splitting of a magnetically sensitive emission line. The first order Zeeman effect yields the longitudinal magnetic-field component, integrated along the line of sight. This type of measurement is done routinely, nowadays. The magnetic field component transverse to the line of sight is much more difficult to measure, pushing to the limit the currently available technology as well as our ability to treat the radiative-transfer problems involved with the interpretation of raw data. As we find in this meeting, there are several ongoing efforts to measure the transverse magnetic field, exploiting various effects such as the second order Zeeman effect, the Hanle effect and the deduction of the Stokes parameters from observed line profiles. Limitations on available spectral lines with clean profiles and known Zeeman patterns generally limit magnetic field measurements to the photosphere and low chromosphere. There are also the limitations on the spatial and temporal resolutions imposed by the magnetograph. The consequence of these constraints is as follows. Firstly, information on the magnetic fields in the upper reaches of the solar atmosphere must be extrapolated from the magnetic fields measured lower down. For this purpose, suitable theoretical models are needed. Secondly, the present generation of magnetographs have attainable spatial resolutions of not less than a few arc seconds and the observed large scale patterns are relatively steady in time. This implies that large-scale, quasi-static structures are appropriate objects for these kind of studies.

¹ The National Center for Atmospheric Research is sponsored by the National Science Foundation.

The use of magnetograph data to build models is commonly acknowledged to have begun with the suggestion of Schmidt (1964), as a working hypothesis, that the magnetic field above the photosphere is potential, or current-free. Potential fields were studied with enthusiasm for a while. The solar atmosphere is an excellent electrical conductor, of course. Particularly for the active region atmosphere, the large magnetic intensity implies that the electric current must be largely parallel to the magnetic field in order to avoid strong Lorentz forces that cannot be balanced by the relatively weak thermal forces available under quasi-static circumstances (e.g. Gold 1964). Taking the electric current to be exactly parallel to the magnetic field, we have the force-free magnetic field. In general, the force-free field poses a nonlinear intractable mathematical problem. Nakagawa and others in the late sixties suggested, as an improvement over the potential field model, that solar structures might be modeled with the subset of "constant alpha" force-free fields. These fields pose linear solvable boundary value problems, and they are the basis of another phase of activity in magnetic field modeling. The logical further development is to consider nonlinear force-free fields, magnetostatic fields which allow for a static interaction between the magnetic field and plasma, magnetic fields with steady velocity flows and, finally, time-dependent fields, in an ascending order of complexity. Mathematical intractability renders such sophisticated models not practical for phenomenological studies. Generally speaking, they tend to be the pre-occupation of theorists challenged with the task of coming away with a model that is sufficiently simplified so as to be tractable and yet adequate for the purpose of theoretically demonstrating a basic physical point. In the following two sections, I will review briefly previous work on potential and linear force-free fields and discuss some theoretical ideas on nonlinear force-free fields and magnetostatic fields for their possible interests to model builders. Lest I give the impression that the solar atmosphere is a calm static medium, I hasten to emphasize that the study of large-scale, quasi-static structures is a choice dictated by the available temporal and spatial resolutions of existing observing instruments. It is important to caution that the physical picture based on this particular mode of observation may be tempered, perhaps in a radical manner, by small-scale processes, velocity flows or highly time-dependent phenomena which are neglected at this level of consideration. In the last section of the paper, we conclude with a discussion of the primary results presented.

2. Force-Free Magnetic Fields

The force-free approximation neglects all interaction between the magnetic field and the tenuous solar atmosphere. For equilibrium, the electric current must be parallel to the magnetic field, giving rise to zero Lorentz force as expressed by the equation

$$(\nabla \times B) \times B = 0. \quad (1)$$

This equation is often rewritten in the form

$$\nabla \times B = \alpha B, \quad (2)$$

which states that the electric current density $J = c/4\pi \nabla \times B$ is proportional to the magnetic field B . The proportionality factor α is constant along individual field lines,

$$B \cdot \nabla \alpha = 0, \quad (3)$$

as a consequence of Maxwell's equation $\nabla \cdot B = 0$. In general, equations (2) and (3) are nonlinear and it is a difficult question as to what boundary conditions give rise to a meaningful problem to solve for B and α ; see Sturrock and Woodbury (1967) and Low (1982a). Two particular cases lead to solvable problems. The first is the case of $\alpha = 0$ everywhere. There is no electric current in the atmosphere and we have a potential field. The second case is that of an α being uniform in space. Equation (3) is then trivially satisfied. The constant α is a free parameter and equation (2) poses a linear problem for B . Much of previous modeling work, using magnetograph data for inputs, were based on solutions of these two special cases. Let us survey the main results and examine some of the underlying assumptions.

The potential field model was expedient at the time it was proposed (Schmidt 1964). Only the longitudinal magnetic field at the photosphere could be measured with confidence then. If this longitudinal field is taken normal to an infinite plane, these data are all one needs to construct the potential field as a solution to a Neuman boundary value problem. It is customary to also assume that the field at great heights vanishes rapidly in order to complete the specification of the boundary conditions. The problem was extended recently for the case where the longitudinal field is measured away from disk center where it is not normal to the photospheric plane (Semel 1967, Sakurai 1982). In practice, a set of magnetograph data is taken only for a limited area of the Sun. The above planar model is therefore appropriate only for magnetic structures which are wholly contained within the field of view. Otherwise, the potential extrapolation without the input of data outside the field of view is quite meaningless (Levine 1975). Data over the whole disk of the Sun had been used as input in models set in the spherical coordinate system (e.g., Altschuler and Newkirk 1969). For these models, the full disk magnetograph data are taken at meridian crossing over a full rotation. This restricts attention to long-lived structures in the corona. Over the large-spatial scales, comparable to $1R_{\odot}$, being considered, the effect of the solar wind that stretches a portion of the solar magnetic flux into a radial field needs to be incorporated. This is usually accomplished by artificially imposing an outer spherical boundary where the field lines are demanded to be radial. Where this outer boundary is to be located is a free parameter of the model (e.g., Newkirk et al. 1968, Schatten et al. 1969). The above types of boundary conditions requiring no more than the longitudinal magnetic field at the photosphere can also be used to determine the constant α force-free magnetic field as a generalization of the potential field (e.g., Nakagawa et al. 1971, Raadu and Nakagawa 1971, Nakagawa and Raadu 1972, Nakagawa 1973). The constant α is an additional free parameter to control the presence of electric currents in the atmosphere. It was not realized earlier but was demonstrated recently that the boundary value problems for a constant α force-free field may not be unique for a fixed value of α , in that more than one solution can satisfy the same boundary data (Chiu and Hilton 1977).

What have we learned from the models based on potential and constant α force-free fields? A standard procedure in these studies is to plot a selection of the calculated field lines to be viewed projected onto a plane perpendicular to a line of sight. The magnetic field pattern is then compared with plasma structures in the same region observed simultaneously in various emissions. With the assumption that the plasma structures outline individual field lines, an agreement is sought between the two patterns (e.g., Newkirk et al. 1968, Veeder and Zirin 1970, Newkirk and Altschuler 1970, Nakagawa et al. 1971, Frazier 1972, Poletto et al. 1975, Levine 1976). Examples can be found where there is agreement between observed plasma structures and computed potential field patterns. There are also examples where the potential field pattern fails to

agree with the observed plasma structures, whereas the field pattern of a force-free field with a suitable choice of constant α brings about a better agreement. In all cases, the comparison is meaningful only if taken globally. Any attempt to compare detail structures leads to ambiguous conclusions, due partly to the complete freedom to select calculated field lines for comparison and partly to the expectation that a potential, or constant α force-free, assumption can only capture the gross structural properties. These models developed in the past twenty years have been instructive and they provided and continue to provide a basis to study magnetic field morphology. The assumption of a potential or a constant α force-free field was made out of expediency in the absence of reliable transverse field measurements. It may well be a good assumption for a particular given situation. One can conceive of regions on the Sun which happen to be in these special states. However, the criterion to test whether a magnetic-field model fits observation relies on a comparison with observed plasma structures under the assumption that the latter outline the magnetic field lines. This is a questionable assumption. At least, it is not always valid (e.g., Low 1981). The test is thus clouded.

From theoretical considerations, one can argue that the force-free magnetic fields expected to be found in active regions are often the nonlinear type with a spatially-varying α . A magnetic field, initially in a potential state, becomes force-free with flowing electric currents as a result of slow evolution driven by the photospheric convective motions (Sturrock and Woodbury 1967, Anzer 1968, Low 1982a). A key feature of this process is the large electrical conductivity of the solar atmosphere which does not allow the electric currents, induced during the evolution, to dissipate over the timescale of evolution. The current in the force-free field is controlled by the function α . In terms of the magnetic field, α reflects the topology of the twisted magnetic field as it evolves under the "frozen-in" condition. Since the convective motions in the photosphere are plasma-dominated and do not have to correlate with the field configuration above the photosphere, specialized motions are required in order to ensure that the magnetic field evolution avoids nonlinear force-free configurations. In general, then, one can argue for the solar magnetic field to be force-free with a non-constant α (Sakurai 1979; Low 1982a). To allow for the possibility that the magnetic field may be a non-constant α force-free field Schmidt (1968) pointed out that the parameter α may be determined by the measurement of all three field components at the photosphere. If we take the photosphere to be the $z = 0$ plane, the z component of equation (2) allows us to compute the distribution of α on the photosphere:

$$\alpha = \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} \right) / B_z \quad (4)$$

It seems clear that the first thing one can do with vector magnetograph data of sufficiently fine spatial resolution is to verify if constant- α force-free fields are realized on the Sun. The matter is a little more complicated. We should take a step back to test if a particular solar magnetic field is indeed force-free before even inquiring if the α of a presumed force-free field is constant. Clearly, with any magnetic field above the photosphere, force-free or not, the vector field measurements yield a photospheric distribution for the right side of equation (4). Whereas, the function α is defined upon the existence of a force-free state. To test whether a magnetic field is force-free, we consider the following well-known property of the Lorentz force.

The Lorentz force can be written as the divergence of the Maxwell stress tensor

$$M_{ij} = -\frac{B^2}{8\pi} \delta_{ij} + B_i B_j / 4\pi. \quad (5)$$

Written this way, the total force in a volume can be converted into the total stress exerted upon the boundary surface (e.g., Parker 1979). To avoid repeating the same qualifying remarks, let us henceforth assume we are considering an isolated magnetic structure located in the infinite half-space above the plane $z = 0$ and vector field measurements, with appropriate spatial resolution for this plane are available. Assuming that the boundaries at infinity do not contribute, that is, the magnetic field in $z > 0$ vanishes at infinity fast enough, the net Lorentz force in the volume $z > 0$ is just the Maxwell stress integrated over the plane $z = 0$. This yields

$$\frac{1}{4\pi} \int_{z=0} B_x B_z dx dy = \frac{1}{4\pi} \int_{z=0} B_y B_z dx dy = \frac{1}{4\pi} \int_{z=0} (B_x^2 - B_z^2 - B_y^2) dx dy = 0 \quad (6)$$

for the 3 components of the total force to vanish if the magnetic field is force-free. Therefore, if all three components of the magnetic field at the photosphere can be measured reliably, equation (6) provides a test for whether the force-free state exists. It should be emphasized that equation (6) is only a necessary condition since some information is lost in the integration to produce equation (6). If we can meet this necessary condition, we may hypothesize that the field is force-free, and equation (4) is then meaningful, yielding the distribution of α on the photosphere. To handle the exact relations in equation (6) in a practical test, we need a criterion for satisfaction of the equations. To that end, we can construct the total magnetic pressure $B^2/8\pi$ from the three measured magnetic field components. The force due to the distribution of the magnetic pressure on $z = 0$, namely,

$$F_0 = \frac{1}{8\pi} \int_{z=0} (B_x^2 + B_y^2 + B_z^2) dx dy \quad (7)$$

is the characteristic magnitude of the total Lorentz force that can be brought to bear on the atmosphere if the magnetic field is not force-free. A practical criterion for satisfying equations (6) is that each of the integrals in equation (6), normalized with respect to F_0 , is much less than unity. Deep in the photosphere, the plasma beta is large and the field above that particular layer includes a region where a substantial Lorentz force may interact with the photospheric plasma. Thus, vector field measurements for this deep layer are not likely to satisfy equations (6). For field measurements in the higher layers, say, the chromosphere, we expect the magnetic field, if it is very intense, to be force-free and the field measurements may satisfy equations (6). These simple tests will make interesting studies.

If the quality of vector field measurements is good enough to allow the verification of equation (6) for a force-free state and, with equation (4), the distribution of α is also obtained, it remains a difficult task to extrapolate for the force-free field above $z=0$, except when α is fortuitously a constant. A mathematical theory for such a nonlinear extrapolation is not available, as far as I am aware (but see Sakurai 1981). One promising method is the use of variational formulation of the problem (Sakurai 1979). While we urge theorists to overcome this technical difficulty, the measurement of the boundary vector field is actually quite sufficient for the determination of the energy of the

magnetic field if it can be first tested by equation (6) to be force-free (Low 1982b). The free energy of the force-free field can also be derived directly from the data. This quantity, which is most important in flare study, can be obtained in spite of our inability to extrapolate for the nonlinear force-free magnetic field in the region $z > 0$.

Taking the scalar product of the Lorentz force with the position vector $\underline{r}(x, y, z)$, integrating over the volume and applying suitable vector identities, one obtains for a force-free field,

$$\int_{z>0} \frac{B^2}{8\pi} dx dy dz = \frac{1}{4\pi} \int_{z=0} (xB_x + yB_y) B_z dx dy, \quad (8)$$

giving the total magnetic energy in $z > 0$ in terms of the three components of the field at $z = 0$. This virial relationship is well-known (e.g., Chandrasekhar 1961, Molodensky 1974). What appears not to have been noticed is its obvious application to obtain the total energy of a force-free field solely in terms of the data provided by vector magnetographs (Low 1982b). Given the measured normal component B_z at $z = 0$, these data determine a unique potential field B_p . If the actual magnetic field in the solar atmosphere is potential, the measured transverse field at $z = 0$ agrees everywhere with the transverse components of the computed B_p . Otherwise, the two transverse fields at $z = 0$ do not coincide. In that case, we can calculate the energy contained in the potential field B_p , which is also force-free, using equation (8),

$$\int_{z>0} B_p^2 / 8\pi dx dy dz = \frac{1}{4\pi} \int_{z=0} (xB_{pz} + yB_{py}) B_z dx dy \quad (9)$$

The difference between the two energies in equations (8) and (9) gives the free energy of the force-free field B in the usual definition (e.g., Low and Hu 1983). It is the free magnetic energy, and not the total magnetic energy, that is believed to fuel the flare. Low (1982b) showed a theoretical example of an evolving force-free field whose total energy is constant in time whereas the free energy increases with time. It will be worthwhile to use magnetograph data to study the change of this free energy during the evolution of a real magnetic structure, in particular, in relation to flares.

To conclude this section, let us say something about the intrinsic information content of a set of vector field measurements taken at the photosphere, quite independent of whether the magnetic field above is force-free (Hagyard et al. 1981). The measurements determine the potential field B_p defined above. Let us define the residual field $B_c = B - B_p$ where B represents the true magnetic field. Since all normal flux B_z at $z = 0$ have been accounted for by B_p , B_c has zero flux everywhere across $z = 0$, being generated by the electric currents in $z > 0$ combined by their image currents in $z < 0$. Clearly, if there is no electric current in $z > 0$, $B_c = 0$ and $B = B_p$. To determine B hinges on how well we can determine B_c . Information of its presence is contained in the distribution $B_{cx} = B_x - B_{px}$, $B_{cy} = B_y - B_{py}$ at $z = 0$, B_{cz} being zero at $z = 0$. This information is all there is on the electric current in $z > 0$, and is inadequate to determine the distribution of B_c in $z > 0$. Only certain topological characteristics of the electric currents in $z > 0$ which give rise to a non-vanishing B_c can be deduced from B_{cx} and B_{cy} at $z = 0$ (see Hagyard et al. 1981). If equation (6) can be used to verify that B is consistent with the force-free state, one can devise methods to go a step further to construct the force-free field in $z > 0$ from the boundary values at $z = 0$. Otherwise, there is a net Lorentz force and no extrapolation is possible without additional information on the plasma distribution that reflects the interaction between the magnetic field

and plasma. Even when equation (6) is satisfied, it is only a necessary but not sufficient condition for a force-free state to obtain. The extrapolated force-free field needs to be subject to further tests, yet to be devised, to verify it represents the true magnetic field.

3. Magnetostatic Fields

The static equilibrium of a magnetic field in the presence of both pressure gradient and gravity is described by the equation

$$\frac{1}{4\pi}(\nabla \times H) \times H - \nabla p - \rho g \hat{z} = 0, \quad (9)$$

where, for simplicity, we have assumed a uniform gravitational acceleration g . This equation, involving the plasma distribution, is more complex than the force-free equation (1) and it is not possible to treat it in a general way. In this section, I will describe briefly some major concerns and results from basic theory and then point out some recent developments in model building.

The basic structure of equation (9) is the following. Since the Lorentz force is always perpendicular to the magnetic field, the weight of the plasma resolved along the magnetic field must be balanced by pressure gradient alone. Therefore, each magnetic flux tube is an individual hydrostatic atmosphere. The profile of each flux tube atmosphere is determined by the run of temperature along it which depends on the energy transport operating to maintain static equilibrium. Perpendicular to the field lines, we find a Lorentz force and for equilibrium it must be balanced by both pressure gradients and plasma weight, which, at the same time, must preserve their hydrostatic relation along the field lines. If we assume the ideal gas law to relate p and ρ to the temperature T , a specification of the static energy transport together with Maxwell's equation $\nabla \cdot H = 0$ completes the whole set of equations. The equilibrium equation (9) takes a neat form if we assume some symmetry. Consider the case of axisymmetry in the cylindrical coordinate system (Low 1975). To satisfy $\nabla \cdot H = 0$, we represent the magnetic field in the form

$$H = \frac{1}{R} \left[\frac{\partial H}{\partial z} \hat{R} - \frac{\partial H}{\partial R} \hat{z} + K \hat{\phi} \right]. \quad (10)$$

Then, equation (9) reduces to the two equations

$$R \frac{\partial}{\partial R} \left[\frac{1}{R} \frac{\partial H}{\partial R} \right] + \frac{\partial^2 H}{\partial z^2} + K \frac{dK}{dH} + 4\pi R^2 \frac{\partial p(z, H)}{\partial H} = 0, \quad (11)$$

$$\frac{\partial p(z, H)}{\partial z} + \rho g = 0, \quad (12)$$

where, as part of the equilibrium requirement, the azimuthal magnetic field has been constrained so that K is a strict function of H and we have, where indicated, chosen the independent variables to be z and H instead of z and R . Equation (12) is the hydrostatic relationship along the field lines and equation (11) is essentially, the force balance perpendicular to the field lines. In principle, there are two independent directions perpendicular to a given direction. It is the assumption of axisymmetry (i.e., $\partial/\partial\phi \equiv 0$) that relates the otherwise two independent directions so that the perpendicular force balance is described by a single equation. Equilibrium solutions to equations such as

those describing the axisymmetric system have been explored as far as these problems are tractable to understand the basic properties of magnetostatic equilibria (e.g. Dungey 1953, Lerche and Low 1980, Low 1981, Hundhausen, Hundhausen and Zweibel 1981). The availability of exact equilibrium solutions also allowed the study of linear stability (e.g., Migliuolo 1982, Zweibel 1982, Low 1982c).

All the above work on equilibrium states is based on the mathematical idealization of some symmetry, such as axisymmetry ($\partial/\partial\phi=0$) or invariance in some given direction ($\partial/\partial x=0$). To take a step closer to geometrically realistic configurations, one would like to give up that idealization. Here, theorists in recent years have found this development very interesting. To give up the simplification associated with symmetry is not merely adding a degree of complexity to the mathematical problem. It raises a non-trivial question of whether an equilibrium is admissible at all, for a given circumstance. In the study of astrophysical magnetic fields, this remarkable property was first pointed out by Parker (1972, 1977a, 1979, 1981a, b, 1983) in his study of nonequilibrium (see also Yu 1973, Rosner and Knobloch 1982, Tsinganos 1982b, Tsinganos et al. 1984). It is conceptually helpful to distinguish between instability and nonequilibrium. Instability refers to an existing equilibrium having the property that a slight imbalance of forces can, in the resulting dynamics, take the system away from that equilibrium. Nonequilibrium is a far more drastic situation where certain constraints of a system preclude a force balance altogether. In astrophysical hydromagnetics, the magnetic Reynolds number is so large that the "frozen-in" approximation applies. The analyses of Parker showed that the frozen-in condition combined with a lack of invariance is a severe constraint that leads, generally, to nonequilibrium. Nature's way out is often to develop local steepening of magnetic field gradients to form electric current sheets. In these localities, dissipation becomes important, the frozen-in condition breaks down and the equilibrium-forbidding constraint is thus removed to allow the system to seek a relaxed state.

The role of symmetry in the balance of forces is implicit in the above derivation of equation (11). There, axisymmetry couples the force balances along the two independent directions perpendicular to a field line so that only one equation arises. Without such a symmetry, we would have obtained two independent statements for the balance of forces in the two independent perpendicular directions. From these two independent statements, the plasma distribution can be eliminated to reduce to a single compatibility relation applied to the magnetic field alone. This relation singles out those magnetic field geometries which are compatible with equilibrium. The theory can be given a precise formulation in terms of the Euler potentials (Low 1980a, Rosner, Low and Holzer 1984). Two implications follow from the formulation. Firstly, the compatibility relation is trivially satisfied for symmetric systems with an ignorable coordinate, the kind of symmetric systems considered in previous work on magnetostatic equilibrium. It is for this reason that the compatibility relation did not arise in previous work. Secondly, the compatibility relation can be treated directly to discover all the admissible magnetic fields which are compatible with equilibrium. Work was initiated recently in this direction and it was found that large classes of three-dimensional equilibrium states exist for which no ignorable coordinates can be identified. These magnetic fields are symmetric in a certain generalized sense in that regularity has been imposed upon the equilibrium magnetic surfaces, through the compatibility relation. For the first time, three-dimensional equilibrium states with rather realistic field geometry become available for theoretical analyses (Hu, Hu and Low 1983, Hu 1984, Low 1982c, 1984, Low, Hundhausen and Hu 1985). Figure 1 shows a three-dimensional, isothermal, magnetized, inverted-U plasma loop embedded

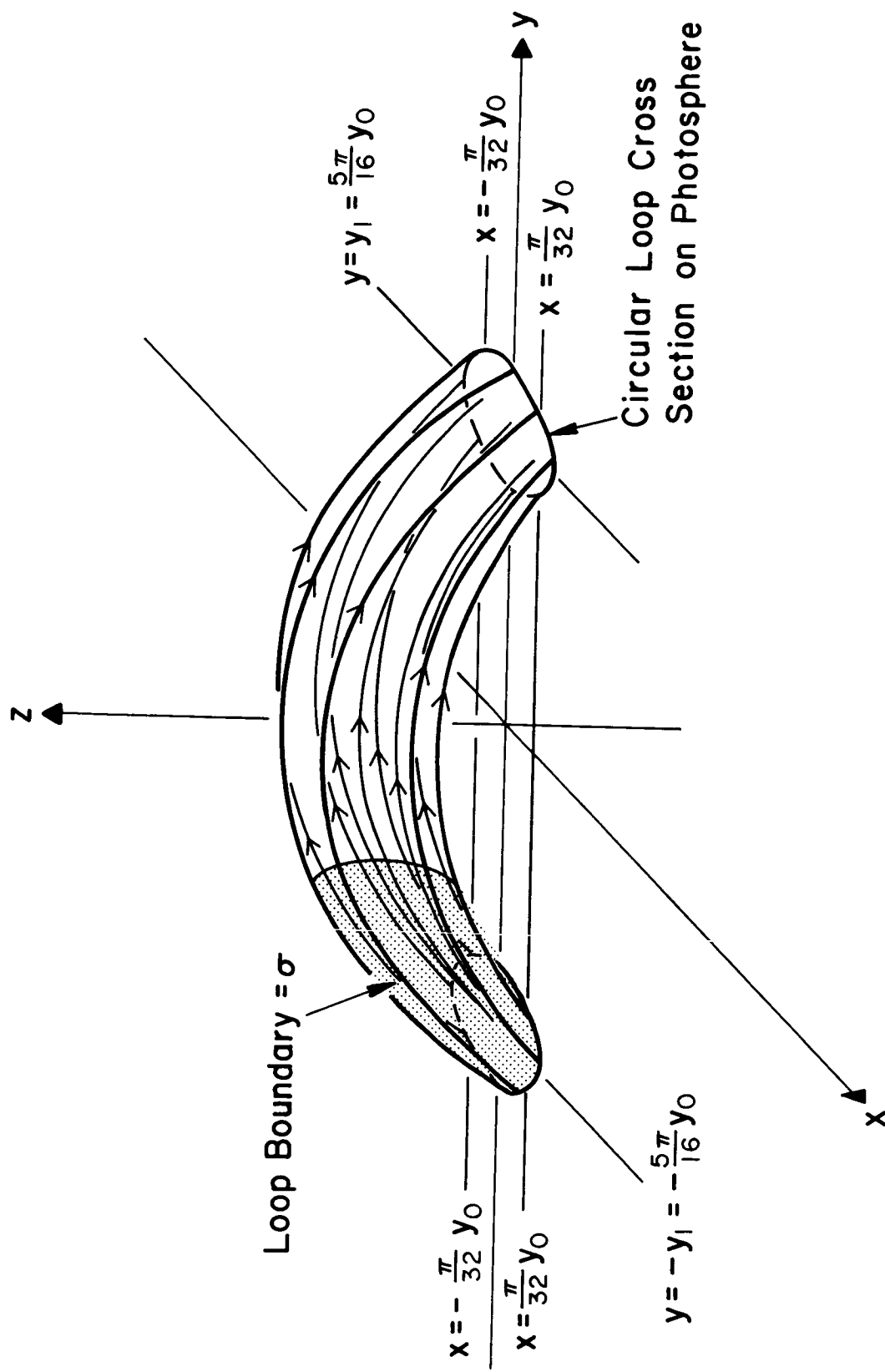


Figure 1. A three-dimensional bipolar plasma loop with its ends rooted to the plane $z = 0$. The arrowed curves are magnetic field lines drawn in perspective. The shaded surface represents a part of the loop boundary σ .

in an unmagnetized isothermal hydrostatic atmosphere of the same temperature and it is an example of the new class of magnetostatic solutions reported in Low (1982c). This particular example shows the plasma loop to have lower density than its surrounding. It is thus bouyant in the stratified atmosphere but is anchored to the lower boundary, in equilibrium, by magnetic tension forces. This is a highly idealized model. In the realistic solar atmosphere, we expect the plasma loops to be embedded in a global continuous magnetic field, such as may be modeled with more sophisticated constructions outlined in Low (1982c). It is a common practice, in discussing observation, to identify observed plasma loops with discrete, imagined, magnetic flux ropes. I believe this is an oversimplification and the study of three-dimensional magnetostatic states will prove helpful in acquainting ourselves with the kind of magnetic field-plasma morphologies which are physically admissible. The solution in Figure 1 is analytic and its magnetohydrodynamic stability can be analyzed by exact methods to show that it is linearly stable under isothermal conditions. Although we have been discussing plasma loops on the Sun for decades, this is the first theoretical three-dimensional demonstration of such an equilibrium state in a stratified atmosphere. It is also one of few known stable equilibria.

It is clear that the complexity of magnetostatic equilibrium does not allow detailed, quantitative modeling. The main viable approach will, for a long time, remain the synthesis of properties we learn theoretically from particular solutions with all available observational data, both from magnetographs as well as from instruments operating in the various spectral ranges to probe the plasma state. I will mention some recent progress in this kind of work in the study of sunspots and prominences.

To the extent that some observed isolated sunspots can be approximated by an axisymmetric system, solutions to equations (11) and (12) provide magnetostatic models to compare with observation. This kind of work dates back to 1950's when Schlüter and Temesvary (1958) proposed their well-known self-similar sunspot model. Unable to treat the full magnetostatic problem, they showed that an exact relationship between the profiles of the atmospheres along the sunspot axis and along a vertical far away from the sunspot can be worked out. Following the derivation of equation (11) and (12), Low (1980b) extended the work of Schlüter and Temesvary to describe the magnetostatic atmosphere of the self-similar sunspot at all points in space. The model showed that the cool interior does not extend more than a few scale heights below the photosphere and that the sunspot magnetic field fans rapidly to approach being nearly potential at the photospheric level so that its Lorentz force exerted on the photosphere is less than what the magnetic pressure alone would suggest. The construction of magnetostatic models can exploit the following technical facility. Equations (11) and (12) contains three unknowns, H , p and ρ ; for simplicity, let us regard $K(H)$ to be a known function. The static energy equation is the additional equation needed to close the system. In general, such a complete set of nonlinear equations is difficult to treat and there is also uncertainty, at the present, on the energy processes operating in the sunspot. For example, as to what cools the sunspots is a controversial question (e.g., Cowling 1976, Giovanelli et al. 1978, Parker 1977b, 1978, Spruit 1977). One simple approach to model building is to ignore the energy equation first, and prescribe a magnetic field by giving the function H , based on various physical considerations. To specify the Schlüter-Temesvary self-similar sunspot magnetic field is an example of this procedure. Then, equations (11) and (12) serve to determine the pressure and density distributions in equilibrium with the prescribed magnetic field (Low 1980b). In a final step, an equation of state like the ideal gas law defines the temperature. The goal of the

modeling is to produce a magnetostatic state that has properties as close to those of a sunspot as possible. In this way, a variety of separate observational information may be synthesized with the model into a more complete picture than can be obtained from the separate observations. However, a word of caution should be added. The mathematical mapping from a prescribed magnetic field to the temperature distribution, in the above model construction, is highly nonlinear. A slight change in field geometry can lead to a rather drastic change in the temperature distribution and it is not easy to see the physical significance of this without considering the energy transport explicitly.

The Schlüter-Temesvary sunspot is a monopole flux that fans outward with all flux going to infinity. Osherovich (1982) proposed, as an improvement, that an isolated sunspot may have only a fraction of its magnetic flux extended to infinity, with the remaining flux re-entering the photosphere in the immediate surrounding of the sunspot. By choosing a suitable function $H(r, z)$ to generate the so-called "return-flux" magnetic geometry, equations (11) and (12) can be solved to compute the equilibrium plasma distribution. The implication of the magnetic geometry for the structure of the atmospheric plasma has been explored in a recent model synthesized with observational inputs (Flå et al. 1982). The model relates the magnetic field gradient on the spot axis with the maximum magnetic field strength and the Wilson depression phenomena. It also suggests a magnetic origin for the umbra-penumbra and penumbra-quiet Sun boundaries. Departure from axisymmetry has been treated. Hu (1984) modeled the radial dark filamentary fibrils that make up the penumbra of a sunspot and suggested that they are bundles of magnetic flux of tubes of lower temperature with a weak density change relative to the bright background. Models for a pair or more sunspots have just been developed (Low 1985). It is clear that static equilibrium is an oversimplification of the real sunspot with its many dynamical attributes, such as the Evershed flows. A magnetostatic model is only a beginning step towards a more quantitative consideration, and away from several decades of models based on qualitative sketches.

Although the prominence phenomenon has been studied with fascination for so long, we still do not know the three-dimensional extent of the prominence as a magnetic object. The difficulty lies partly in our inability to observe the global exterior coronal magnetic field supporting the prominence and partly in the intractability of theoretical models needed to provide insight. Through a variety of two- and three-dimensional magnetostatic solutions, the following structural properties seem to have emerged. In the theoretical attempt to construct high density enhancements to correspond to H_α prominence matter, the solution also readily manifests density depletion regions. Perhaps we have been overlooking the importance of density depletions in the structure of prominences. Any condensation of prominence matter must lead to the Parker instability first discussed for the interstellar clouds (Parker 1966, Sweet 1971). The dense condensate slides along magnetic field lines to the lowest points or valleys. The evacuated magnetic crest, being buoyant, would rise. An added effect is that the distorted magnetic field enhances its effect as a thermal shield for the condensate, which probably, in the first place, is responsible for the thermal instability that starts the whole process. The situation runs away until an equilibrium is reached where the buoyancy of density depleted regions are pegged by downward acting magnetic tension forces and the heavy condensates are supported by local upward-acting magnetic tension forces. The prominence high density plasma is thus very inhomogeneous and lumpy, interspersed between density depletions. The buoyancy of the magnetic structure due to the density depletions plays an

important role in the eruption of the prominence, a point appreciated only recently in relation to coronal transients (Fisher and Poland 1981, Low, Munro and Fisher 1982). It is interesting to note once the global prominence has been created, the depletion regions are trapped in it, whereas density enhancements can be subject to leaking to reduce the total weight. This tendency works towards the entire structure becoming buoyant. Finally, in contrast to models based on symmetric equilibrium states, three-dimensional models readily exhibit electric currents flowing into and out of the dense regions of the prominence. Where the density is low, the electric currents are maintained in force-free states.

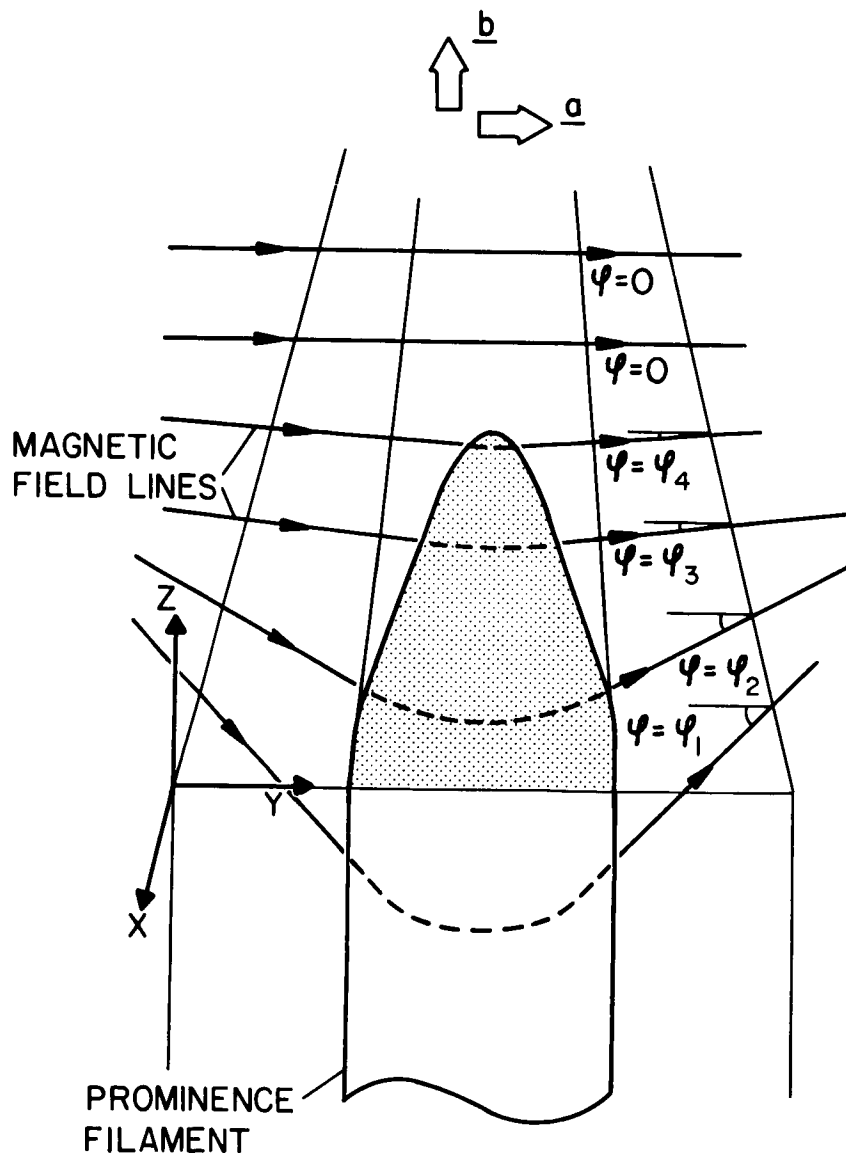


Figure 2. A prominence filamentary structure suspended vertically by magnetic field lines. Different properties in the directions \underline{a} and \underline{b} give rise to anisotropic structures.

On the small scale, the magnetic field gives the embedding plasma a highly anisotropic structure, as illustrated in Figure 2 showing a clump of plasma suspended by magnetic field lines. In the direction \mathbf{a} , the transition from high density in the plasma clump to the low density is essentially controlled by the scale height arising from the strictly hydrostatic relation along the magnetic field. The transition from the low temperature inside the prominence plasma to the exterior high temperature is controlled by thermal conduction along the magnetic field. It is readily shown that both these transitions are diffuse and cannot be less than 300 km thick. On the other hand, quite abrupt jumps in density and temperature are possible in the direction \mathbf{b} , perpendicular to the magnetic field. In that direction, large surface electric currents can confine the plasma with a sharp boundary while thermal conduction is quite negligible. The above and other properties await further studies with observations to be carried out to fine spatial resolutions.

The prominence model described above has a global magnetic configuration first envisioned by Kippenhahn and Schlüter (1957). A natural consequence of the condensation discussed is that the prominence tends to lie parallel to the photospheric magnetic polarity inversion line and that the magnetic polarity of the prominence, as fixed by the sense of the threading magnetic field lines, is the same as that of the photospheric polarity. An alternative mode of prominence formation is that of Kuperus and Raadu (1974), where reconnection of open magnetic fields leads to condensations having a magnetic polarity opposite to that of the photosphere. As to which of these two contrasting configurations takes place on the Sun, it became an interesting issue recently when results of vector measurements of prominence magnetic fields were made available. Leroy et al. (1983), based on a rather large sample of crown filaments, concluded that there is a statistical preference for the Kuperus-Raadu configuration. Whereas, Athay et al. (1983), based on a smaller sample of equatorial prominences, concluded that the observation is consistent with both Kuperus-Raadu and Kippenhahn-Schlüter configurations depending on the interpretation made additionally on the inherent ambiguity in the sense of the measured transverse-field. This development is an exciting contact between models and observational results.

IV. Discussion

I have concentrated on theoretical results in the study of force-free magnetic fields and magnetostatic fields. Although the solar atmosphere is evidently dynamic, with all kinds of flows and explosive behaviors, the assumption of a static state is appropriate when considering large-scale structures and quasi-static evolution and this is a regime observable with the usual magnetographs. In any case, recent theoretical progress in the study of force-free and magnetostatic fields has provided a firm basis for an understanding of the physical properties of these magnetic fields. We are thus in a position to embark on quantitative modeling with observational data, in the manner discussed in the previous two sections. Such an effort is a first step towards more realistic quantitative considerations which will include velocity flows and other dynamical effects that have been neglected. Dynamical effects are understandably much more complex and less accessible to exact theoretical treatment and we refer the readers elsewhere for reviews, of recent developments (e.g., Parker 1979, 1982a, b, c, d, Tsinganos 1981, 1982a, Spruit 1983).

In section II, we advocated going beyond potential and linear force-free field models to test for the existence of force-free states by the use of vector

magnetograph data with classical virial relations. This is a physically well-defined project with little *ad hoc* assumptions. When we can, indeed, locate a force-free field on the Sun, it is then interesting to determine whether it is of the linear or nonlinear type by deriving the α -distribution directly from the magnetograph data. Extrapolation of nonlinear force-free fields above the photosphere remains an intractable computational problem and it will be worthwhile to devise suitable methods of solution. On the other hand, without knowing the precise magnetic field distribution, we can derive its total magnetic energy and free magnetic energy, quantities of great interest to the study of flares, making use of vector magnetograph data with a certain well-known virial relation. The observational application of this particular virial relation was appreciated only recently.

Models of magnetic fields in static equilibrium with pressure gradients and gravity are difficult to build. In section III, we briefly reported the progress of constructing particular mathematical solutions from the consideration of two-dimensional, or, symmetric systems to that of truly three dimensional systems which one can hope to compare with realistic structures. The addition of a third spatial dimension to the equilibrium problem changes the problem in a physically profound way, a point first realized by Parker. Once this basic property of magnetostatic equilibrium is recognized, it becomes straightforward, conceptually, to devise methods of building three-dimensional magnetostatic configurations. This kind of work is at an early development but we already have, for the first time, simple theoretical models for such objects as the three-dimensional ubiquitous inverted-U plasma loops that populate the solar atmosphere. Detailed quantitative modeling with observational data will not be possible partly because the magnetostatic equations cannot be treated in a general way and partly because of our inability to measure the magnetic field in the upper reaches of the solar atmosphere. A synthesis between our theoretical understanding of the magnetostatic morphology and physical properties, and various available observations will remain the obvious viable approach, such as illustrated by the recent modeling of sunspots.

We should always caution against the oversimplification of the real phenomena when seen through these static models. The actual modeling studies may also run into difficulties with instrumental spatial resolutions, problems arising from boundary conditions for magnetic structures which are not spatially isolated, etc. However, the case can be argued for genuine prospects of learning considerable physics. This is an incentive for us to intensify the efforts to improve magnetograph technology and to solve the difficult radiative-transfer problems encountered in the interpretation of magnetograph raw data.

I thank Grant Athay for helpful comments.

References

- Altschuler, M. D. and Newkirk, G., Jr. 1969, *Solar Phys.* **9**, 131.
- Anzer, U. 1968, *Solar Phys.* **3**, 298.
- Athay, G., Querfeld, C., Smartt, R.N, Landi Degli-Innocenti, E. and Bommier, V. 1983, *Solar Phys.* **89**, 3.
- Babcock, H. W. and Babcock, H. D. 1955, *Ap. J.*, **121**, 349.
- Chandrasekhar, S. 1961, *Hydrodynamic and Hydromagnetic Stability*, Oxford Univ. Press.
- Chiu, Y. T. and Hilton, H. H. 1977, *Ap. J.*, **212**, 873.
- Cowling, T. G. 1976, *M.N.R.A.S.*, **177**, 409.
- Dungey, J. W. 1953, *M. N. R. A. S.* **113**, 180.
- Fisher, R. and Poland, A. I. 1981, *Ap. J.*, **246**, 1004.
- Flå, T., Osherovich, V. A. and Skumanich, A. 1982, *Ap. J.*, **261**, 700.
- Frazier, E., N. 1972, *Solar Phys.* **24**, 98.
- Gionvanelli, R. G., Livingston, W. C. and Harvey, J. W. 1978, *Solar Phys.*, **59**, 49.
- Gold, T. 1964, in W. N. Hess (ed.) *AAS - NASA Sympo. Physics of Solar Flares*, NASA, 389.
- Hagyard, M. Low, B. C. and Tandberg-Hanssen, E. 1981, *Solar Phys.*, **73**, 257.
- Hu, Y. Q. 1984, *Chinese Astron. Astrophys.* **8**, 41.
- Hu, W. R., Hu, Y. Q. and Low, B. C. 1983, *Solar Phys.*, **83**, 195.
- Hundhausen, J. R., Hundhausen, A. J. and Zweibel, E. G. 1981, *J. Geophys. Res.*, **86**, 11,117.
- Kippenhahn, R. and Schlüter, A. 1957, *Zs. Ap.*, **43**, 36.
- Kuperus, M. and Raadu, M. A. 1974, *Astron. Astrophys.*, **31**, 189.
- Lerche, I. and Low, B. C. 1980, *Solar Phys.*, **67**, 229.
- Leroy, J. L., Bommier, V. and Sahal-Brechot, S. 1983, *Solar Phys.*, **83**, 135.
- Levine, R. H. 1975, *Solar Phys.*, **44**, 365.
- Levine, R. H. 1976, *Solar Phys.*, **46**, 159.
- Low, B. C. 1975, *Ap. J.*, **197**, 251.
- Low, B. C. 1980a, *Solar Phys.*, **65**, 147.
- Low, B. C. 1980b, *Solar Phys.*, **67**, 57.
- Low, B. C. 1981, *Ap. J.*, **246**, 538.
- Low, B. C. 1982a, *Rev. Geophys. Space Phys.*, **20**, 145.
- Low, B. C. 1982b, *Solar Phys.*, **77**, 43.
- Low, B. C. 1982c, *Ap. J.*, **263**, 952.
- Low, B. C. 1982d, *Solar Phys.*, **75**, 119.
- Low, B. C. 1984, *Ap. J.*, **277**, 415.
- Low, B. C. 1985, *Ap. J.*, to be submitted.
- Low, B. C. and Hu, Y. Q. 1983, *Solar Phys.*, **84**, 83.
- Low, B. C. Hundhausen, A. J. and Hu, Y. Q. 1985, *J. Geophys. R.*, to be submitted.
- Low, B. C., Munro, R. and Fisher, R. 1982, *Ap. J.*, **254**, 335.

- Migliuolo, S. 1982, *J. Geophys. Res.*, **87**, 8057.
- Molodensky, M. M. 1974, *Solar Phys.*, **39**, 393.
- Nakagawa, Y. 1973, *Astron. Astrophys.*, **27**, 95.
- Nakagawa, Y. and Raadu, M. A., 1972, *Solar Phys.*, **25**, 127.
- Nakagawa, Y., Raadu, M. A., Billings, D. F. and McNamara, D. 1971, *Solar Phys.*, **19**, 72.
- Newkirk, Jr., G. and Altschuler, M. D. 1970, *Solar Phys.*, **13**, 131.
- Newkirk, Jr., G. Altschuler, M. D., and Harvey, J. W. 1968, in K. O. Kiepenheuer (ed.) *Structure and Development of Solar Active Regions*, IAU Symp. No. 35, 379.
- Osherovich, V. 1982, *Solar Phys.*, **77**, 63.
- Parker, E. N. 1966, *Ap. J.*, **145**, 811.
- Parker, E. N. 1972, *Ap. J.*, **174**, 499.
- Parker, E. N. 1977a, *Ann. Rev. Astron. Astrophys.*, **15**, 45.
- Parker, E. N. 1977b, *M.N.R.A.S.*, **179**, 94p.
- Parker, E. N. 1978, *Ap. J.*, **221**, 368.
- Parker, E. N. 1979, *Cosmical Magnetic Fields*, Oxford Univ. Press.
- Parker, E. N. 1981a, b, *Ap. J.*, **244**, 631 & 644.
- Parker, E. N. 1982a, b, c, d, *Ap. J.*, **256**, 292, 302, 736, & 746.
- Parker, E. N. 1983, *Ap. J.*, **264**, 635.
- Poletto, G., Vaiana, G. S., Zombeck, M. V., Krieger, A. S. and Timothy, A. F. 1975, *Solar Phys.*, **44**, 83.
- Raadu, M. A. and Nakagawa, Y. 1971, *Solar Phys.*, **20**, 64.
- Rosner, R. and Knobloch, E. 1982, *Ap. J.*, **262**, 349.
- Rosner, R., Low, B. C. and Holzer, T. E. 1984, in P. A. Sturrock, T. E. Holzer and D. Mihalas (eds.) *Physics of the Sun*.
- Rust, D. M. and Roy, J.-R. 1971, in R. Howard (ed.) *Solar Magnetic Fields*, IAU Symp. No. 43, 595.
- Sakurai, T. 1979, *Publ. Astron. Soc. Japan*, **31**, 209.
- Sakurai, T. 1981, *Solar Phys.*, **69**, 343.
- Sakurai, T. 1982, *Solar Phys.*, **76**, 301.
- Schatten, K. H., Wilcox, J. M. and Ness, N. F. 1969, *Solar Phys.*, **6**, 442.
- Schlüter, A. and Temesvary, S. 1958, in B. Lehnert (ed.) *Electromagnetic Phenomena in Cosmical Physics*, IAU Symp. No. 6, 263.
- Schmidt, H. 1964, in W. N. Hess (ed.) *AAS - NASA Symp. Physics of Solar Flares*, NASA, 107.
- Schmidt, H. 1968, in K. O. Kiepenheuer (ed.) *Structure and Development of Solar Active Regions*, IAU Symp. No. 35, 95.
- Semel, M. 1967, *Ann. d'Astrophys.*, **30**, 513.
- Spruit, H. C. 1977, *Solar Phys.*, **55**, 3.
- Spruit, H. C. 1983, in J. Stenflo (ed.), "Solar and Stellar Magnetic Fields: Origin and Coronal Effects", IAU Symp. No. 102, 41.
- Sturrock, P. A. and Woodbury, E. T. 1967, in P. A. Sturrock (ed.), *Plasma Astrophysics*, Academic Press, 155.

- Sweet, P. A. 1971, in R. Howard (ed.), *Solar Magnetic Fields*, IAU Symp. No. 43, 457.
- Tsinganos, K. 1981, *Ap. J.*, **245**, 764.
- Tsinganos, K. 1982a, *Ap. J.*, **252**, 775.
- Tsinganos, K. 1982b, *Ap. J.*, **259**, 832.
- Tsinganos, K., Distler, J. and Rosner, R. 1984, *Ap. J.*, **278**, 409.
- Veeder, G. J. and Zirin, H. 1970, *Solar Phys.*, **12**, 391.
- Yu, G. 1973, *Ap. J.*, **181**, 1003.
- Zweibel, E. G. 1982, *Ap. J. Letters*, **258**, L53.

MAGNETIC AND THERMODYNAMIC STRUCTURE OF THE
SOLAR CORONA DURING SUNSPOT MINIMUM

Vladimir A. Osherovich

Cooperative Institute for Research
in Environmental Science
Colorado University¹

Israel Tzur

National Center for Atmospheric Research²
P. O. Box 3000
Boulder, Colorado 80307

Erast B. Gliner

McDonnell Center for Space Sciences
Washington University
St. Louis, Missouri 63130

Abstract. A model of the solar corona ($1.5R_{\odot} \leq R \leq 5R_{\odot}$) during sunspot minimum is constructed. We suggest that pole-equator asymmetry is caused by interaction of global azimuthal electric currents in the corona and potential magnetic fields originating below the corona. The problem is reduced to a nonlinear ordinary differential equation for a structural function. Taking the difference in electron density between the equatorial plane and the polar direction from observations, we solved the basic equation and derived the global electric current distribution around the Sun along with the magnetic and thermodynamic structures of the solar corona. The asymmetry between the two hemispheres, which can be due to the magnetic quadrupole, is considered in detail. It is shown that an additional quadrupole term in the magnetic field representation affects the position of polar coronal hole boundaries differently in the two hemispheres. As a result, the outflow from the two polar regions is found to be different.

¹Work done at the Space Environment Laboratory, NOAA, ERL, Boulder, Colorado 80303.

²The National Center for Atmospheric Research is sponsored by the National Science Foundation.

1. Introduction

Empirical models of the solar corona are obtained by converting the brightness in eclipse photographs (or photographs obtained with a coronagraph) into electron density N_e . Such modeling initiated by Schuster (1880) and continued by Young (1911), Woltjier (1926), and Minnaert (1930) resulted in the establishment of a spherically symmetric corona model (Baumbach, 1937, 1939). Later, Allen (1946) and van de Hulst (1950) suggested two types of electron density distributions: spherically symmetric distribution for the corona during sunspot maximum, and the distribution during sunspot minimum in which the electron density in the equatorial plane is 2-3 times higher than the density in the polar region. A reliable N_e contributes to our understanding of the coronal phenomenon. However for future studies to find the relation between magnetic, thermodynamic, and dynamic structures we have to consider the balance of forces in the asymmetric solar corona during sunspot minimum. We assume that pole-equator asymmetry in the magnetically dominated region $1.5 \leq r \leq 5$ ($r \equiv R/R_0$),

$$\text{where } \rho \frac{v^2}{2} \ll \frac{B^2}{8\pi},$$

is caused mainly by the Lorentz force. This force is a result of the interaction of the global coronal currents with the underlying potential fields. Thus our approach differs from modeling based on potential or force-free magnetic fields.

2. Quasi-static Approximation

We assume that the magnetic structure of the solar corona during sunspot minimum can be described by a Chandrasekhar potential

$$A = H_0 R_0^2 X_0 [u(r) (1-\mu^2) - X_0^{-1} \mu \operatorname{sign} \mu] \quad (1)$$

where H_0 and X_0 are constants,

u is the structural function,

$\mu \equiv \cos \theta$, θ is a colatitude.

The corresponding magnetic field is

$$\vec{B} = \frac{1}{R_0^2 r} \left(-\frac{1}{r} \frac{\partial A}{\partial \mu}, -\frac{1}{(1-\mu^2)^{1/2}} \frac{\partial A}{\partial r}, 0 \right) \quad (2)$$

As a result we have a combination of a radial field $H_0 r^{-2} \operatorname{sign} \mu \vec{e}_R$ and a dipole-

like field (i.e., a magnetic field with a dipole angular dependence and an arbitrary radial dependence). For this field we solve the magnetohydrostatic problem, neglecting the influence of the outflow on the force balance. Then, using the found magnetic field and the induction equation we obtain the velocity of the outflow that can coexist with such a magnetic structure. The described approach we define as a quasi-static approximation.

The magnetohydrostatic problem for the axially-symmetric magnetic configuration around a gravitating body was reduced by Uchida and Low (1981) to a partial differential equation for gas pressure P and Chandrasekhar's potential A . This equation and a complementary condition allow us to find the density distribution, if $P(r, A)$ is known. Using these results for potential (1) we further reduce the problem to an ordinary differential equation (Osherovich, Tzur and Gliner, 1984):

$$r^2 \left\{ \frac{d}{dr} \left[\frac{1}{r^2} \left(\frac{d^2 u}{dr^2} - \frac{2u}{r^2} \right) \right] \right\} (u + X_0^{-1}) = L \Delta N_e(r), \quad (3)$$

where

$$\Delta N_e(r) \equiv N_e(r, \mu=0) - N_e(r, \mu=1)$$

is the difference in electron density between the equatorial plane and the polar direction, and

$$L \equiv \frac{4\pi \cdot 1.17 M_p g_0 R_0}{H_0^2 X_0^2}, \quad M_p \text{ is a proton mass.}$$

We define the magnetic surface $A(r, \mu)=0$ as a polar hole boundary. This surface separates the polar region with open magnetic lines from the area with magnetic lines that return to the equatorial plane. The parameter X_0 is related to the colatitude of the polar hole boundary at $r=1$, namely

$$X_0 = \frac{\cos \theta_b}{\sin^2 \theta_b}.$$

Taking $\Delta N_e(r)$ from one of the empirical models, and solving the basic equation (3), we find the structural function $u(r)$ that allows us to express magnetic and thermodynamic parameters such as \vec{B} , ρ , P and T analytically. To solve (3) we take the boundary conditions $u(r_0)$ and $u'(r_0)$ (for Allen's model $r_0=1.5$) and then adjust $u''(r_0)$ to obtain a solution with the asymptotic behavior $u \rightarrow r^{-1}$ as $r \rightarrow \infty$. We compare $\Delta N_e(r)$ from Allen's empirical model (1973), from Saito's model (1970), and the suggested Saito-Munro-Jackson model [a model with equatorial N_e from Saito's model and the polar distribution from Munro and Jackson (1977)].

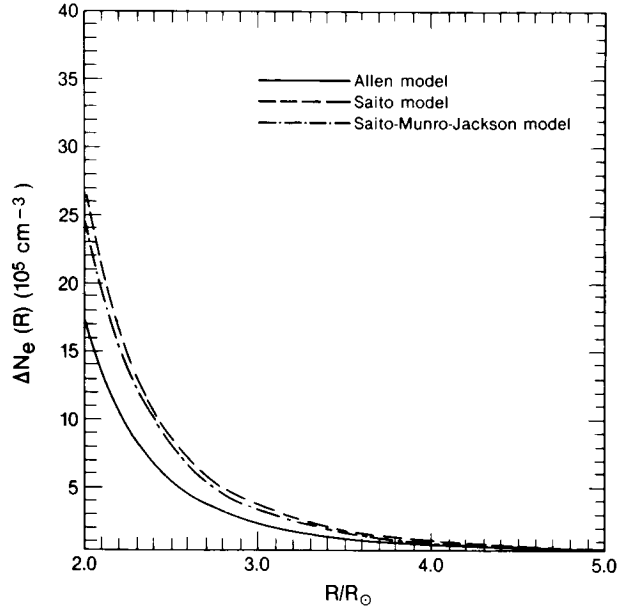


Fig. 1. Difference in electron density between equatorial plane and the polar direction for three empirical models.

Figure 1 shows that curves for $\Delta N_e(r)$ from these three empirical models are similar. Therefore, we present results only for Allen's distribution. We chose $X_0 = \sqrt{2}$ in our calculations, which corresponds to $\theta_p = 45^\circ$. We also chose $H_0 = 2.5G$. The corresponding total magnetic field above the pole $B_R(r=1, \mu=1) = 9.6G$. This value is between the 5G obtained by Stenflo (1971) and 11.5G obtained by the Stanford University group (Hoeksema, Wilcox and Scherrer, 1982).

The coronal magnetic structure for $X_0 = \sqrt{2}$ is presented in Figure 2. The magnetic surface $A(r, \mu) = 0$ corresponds to the boundaries of the polar holes. In the equatorial plane we have a current sheet, due to the discontinuity of the radial field $H_0 r^{-2} \sin \mu \vec{e}_R$. Figure 3 shows the related isodensity contours. Gas pressure and temperature are presented in Figures 4 and 5 for every 15° of the colatitude. From these figures the polar region, as expected, appears to have lower density, pressure, and temperature than the equatorial plane. For the infinitely conductive axisymmetric flow without an azimuthal component ($V_\phi = 0$)

$$V_R = \frac{dF(A)}{dA} \frac{1}{\rho} B_R \quad (4)$$

$$V_{\theta} = \frac{dF(A)}{dA} \frac{1}{\rho} B_{\theta} \quad (5)$$

where $F(A)$ is a stream function.

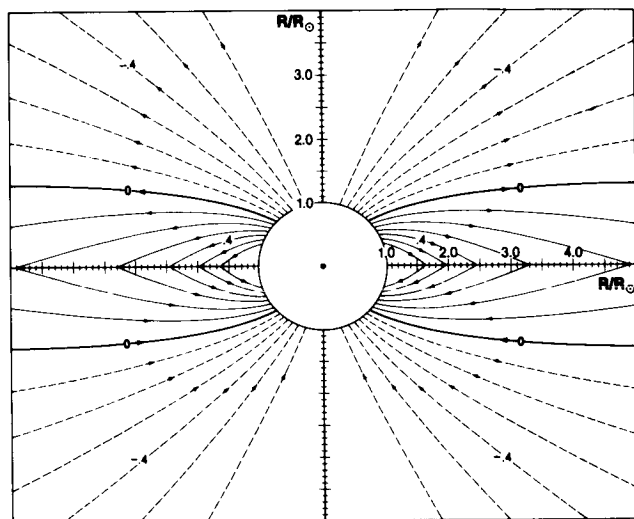


Fig. 2. Magnetic lines of force for the model of solar corona during sunspot minimum without a quadrupole ($X_0 = \sqrt{2}$, numbers on the curves correspond to A/A_0).

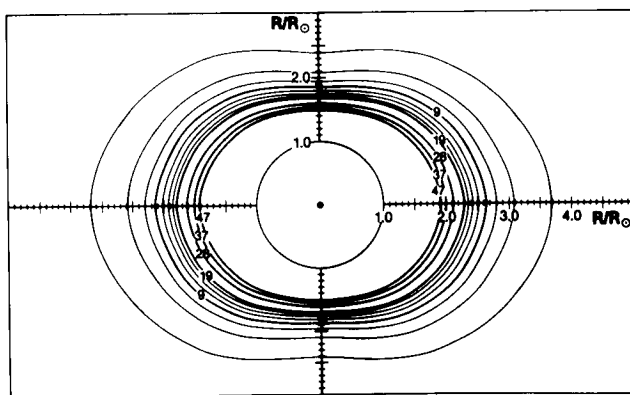


Fig. 3. Isodensity contours for the model of solar corona during sunspot minimum without a quadrupole. (Parameters are chosen the same as for Figure 2). Numbers on the contours correspond to density ($10^{-19} \text{ g cm}^{-3}$).

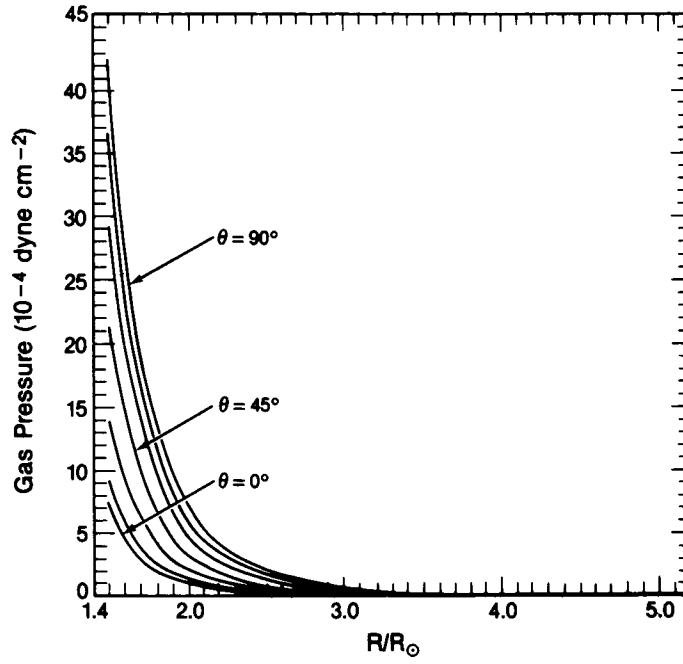


Fig. 4. Gas pressure distribution for the model of solar corona during sunspot minimum [$X_0 = \sqrt{2}$, $\Delta N_e(r)$ is taken from Allen (1973)].

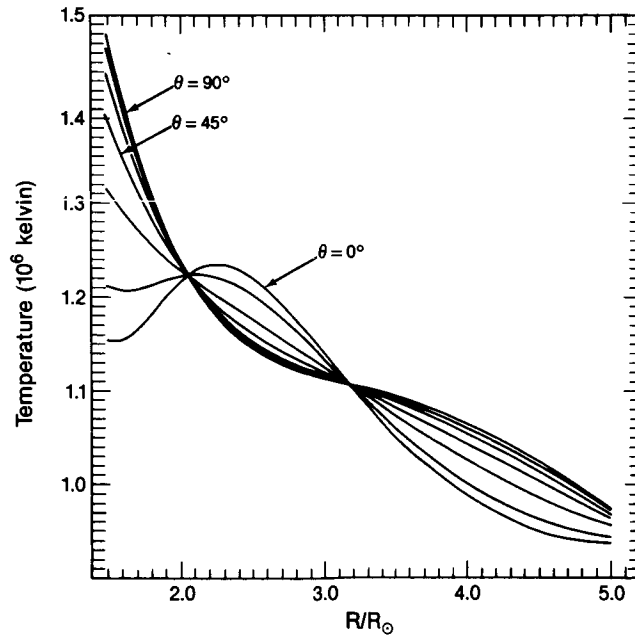


Fig. 5. Temperature distribution for the model of solar corona during sunspot minimum. (Parameters are chosen the same as in Figures 2-4).

We assume $F=F_0A$ in the northern hemisphere and $F=-F_0A$ in the southern hemisphere (Osherovich and Suess, 1982).

Then choosing the constant F_0 in such a way that $V_R (R=1 \text{ A.U.}, \theta=90^\circ) = 400 \text{ km s}^{-1}$ we obtain the radial velocity profiles in the region $1.5 \leq r \leq 5$ (Figure 6).

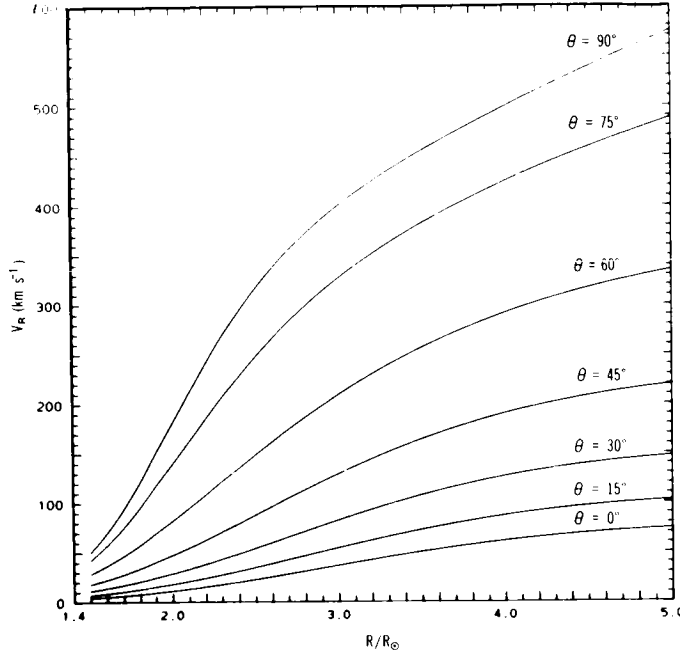


Fig. 6. Radial velocity of the outflow in the corona.

These profiles demonstrate faster acceleration of the solar wind from the polar region than in the equatorial plane.

As long as we consider a combination of a dipole-like magnetic field and the radial field, the corona model is symmetric over the equatorial plane. Introducing a new potential

$$A = H_0 R_0^2 X_0 [u(r) (1-\mu^2) - X_0^{-1} \mu \operatorname{sign} \mu + X_1^{-1} r^{-2} \mu (1-\mu^2)]. \quad (6)$$

With a quadrupole term (the term proportional to X_1^{-1}) we destroy this symmetry. The configuration with the additional quadrupole magnetic field is presented in Figure 7, and the related isodensity contours are shown in Figure 8. Both magnetic and thermodynamic structures now are different for the northern and southern hemispheres: the southern coronal hole has become larger than the northern one. This change influences the outflow from the polar regions (Figure 9). For $r > 3$ the acceleration from a larger polar hole is faster;

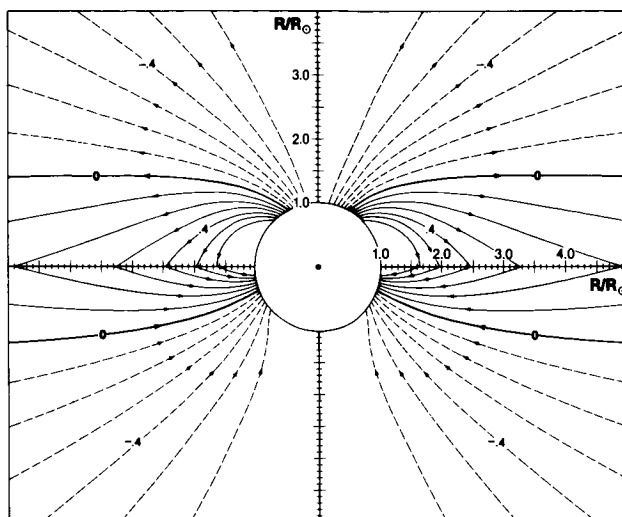


Fig. 7. Magnetic lines of force for the model of solar corona during sunspot minimum with quadrupole ($X_0 = 2$, $X_1^{-1} = \sqrt{2}$).

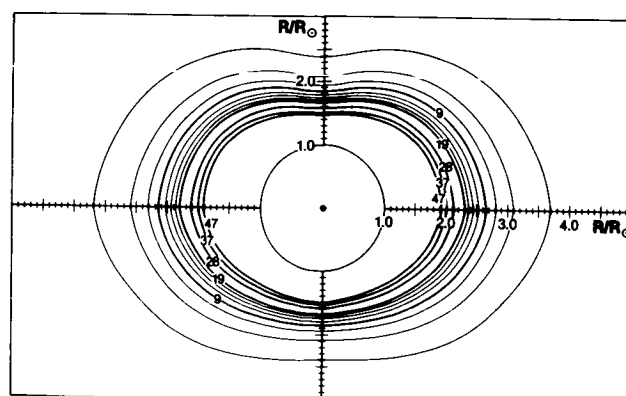


Fig. 8. The influence of a quadrupole on the density distribution in the model of solar corona during sunspot minimum ($X_0 = 2$, $X_1^{-1} = \sqrt{2}$).

$$\left[\text{i.e., } \frac{dV_R(r, \mu=-1)}{dr} > \frac{dV_R(r, \mu=1)}{dr} \right]$$

It was shown earlier (Osherovich et al., 1983) that the outflow increases the temperature as much as 30% compared with the temperature in the quasi-static model. We calculate electric current density j around the Sun, using our quasi-static approximation (Figure 10). To the best of our knowledge this is the first estimation of the global azimuthal current density in the solar atmosphere. The current density ranges from $14 \times 10^{-10} \text{ A m}^{-2}$ ($r=1.5$) to $0.1 \times 10^{-10} \text{ A m}^{-2}$ ($r=5$) and is larger near the equatorial plane, compared with the polar region.

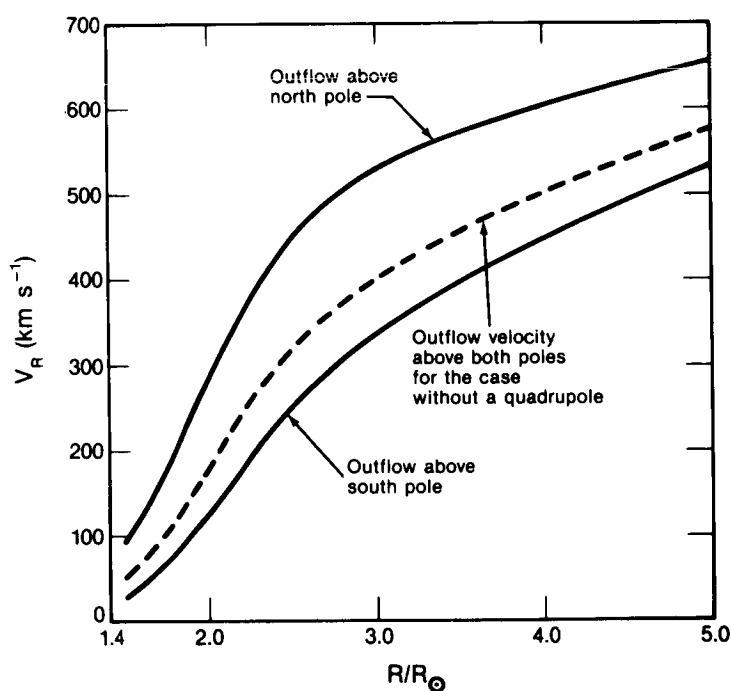


Fig. 9. Influence of the quadrupole magnetic field on the outflow from two polar regions ($X_0 = 2$, $X_1^{-1} = \sqrt{2}$).

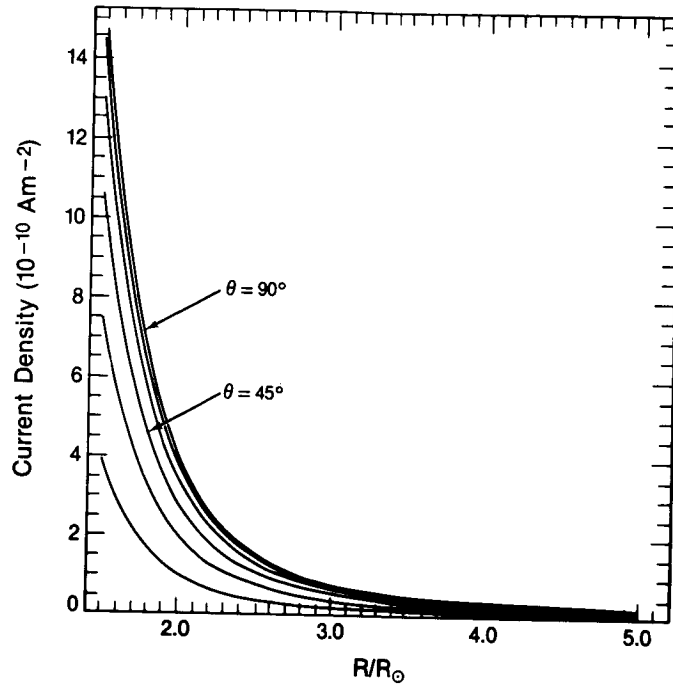


Fig. 10. Electric current density distribution in the solar corona during sunspot minimum. Parameters are the same as in Figures 2-6.

3. Dynamic Approximation

For the quasi-radial flow (V_ϕ , $V_\theta \ll V_R$) it is possible to include the dynamical term in the basic ordinary equation (3). The modified equation for the northern hemisphere is

$$r^2 \left\{ \frac{d}{dr} \left[\frac{1}{r^2} \left(\frac{d^2 u}{dr^2} - \frac{2u}{r^2} \right) \right] \right\} (u + X_o^{-1}) =$$

$$= L \Delta N_e(r) + \frac{4\pi}{H_o^2 X_o^2} r^2 \left[\frac{1}{2} \rho \frac{\partial V_R^2(r, \mu)}{\partial r} \right] \bigg|_{\mu=1}^{\mu=0} \quad (7)$$

where
$$V_R = \frac{F_0}{\rho} B_R.$$

For the dynamic model without a quadrupole ($X_1^{-1}=0$) there is a symmetry between the hemispheres so that the equation for the southern hemisphere coincides with (7). Note that in the dynamic approximation for a model with a quadrupole field equations for the northern and southern hemispheres are different (Osherovich, Gliner and Tzur, 1984).

Equation (7) allows us to study the influence of the outflow on the global electric current around the sun. The current density can be presented in the form

$$j = j_{\text{static}} + j_{\text{dynamic}},$$

where j_{dynamic} represents the contribution to the total current density that is due to the outflow. Figure 11 demonstrates that j_{dynamic} is negative and rather small compared with j in the region $1.5 \leq r \leq 3$. Thus, our quasi-static approximation is relevant in this region. At $r \sim 4$, j_{static} and j_{dynamic} are comparable.

It is clear that our dynamic model is applicable also for $r > 5$. The farther we go from the Sun, the better the quasi-radial approximation.

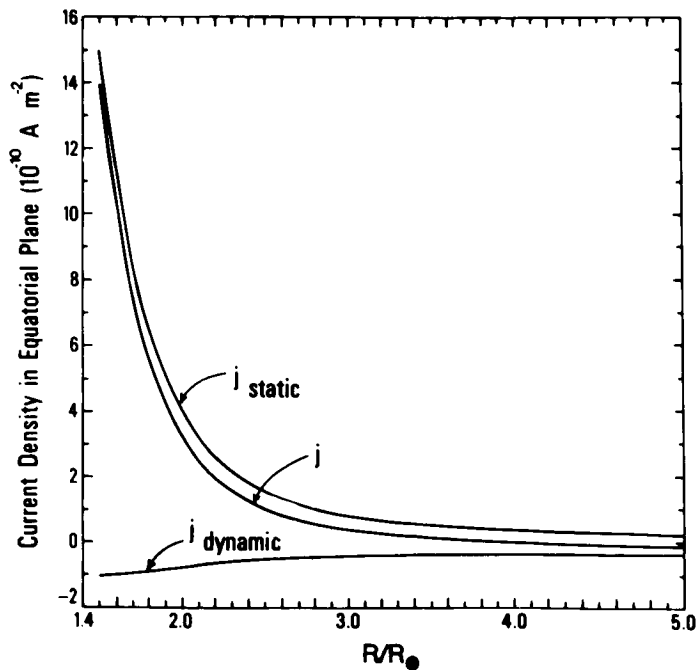


Fig. 11. Comparison of the electric current densities in the equatorial plane for the dynamic model (j) and the quasi-static model (j_{static}); j_{dynamic} represents the contribution of the dynamic term to the total current density j .

4. Discussion

The suggested model has few drawbacks. The current sheet in the equatorial plane is an oversimplification of the real magnetic structure in the vicinity of the equatorial plane. However, a more advanced model with continuous magnetic field in the equatorial plane has already been suggested [Osherovich, Tzur and Gliner, 1984 (Appendix)]. This model includes a radial field with arbitrary latitudinal dependence.

Our present model treats pole-equator asymmetry and the asymmetry between the hemispheres (e.g., as observed by Moussas et al., 1983). However, it is still an axisymmetric model. Longitudinal dependence should be incorporated in the model. In spite of these disadvantages, the model provides better understanding of the relation between magnetic, thermodynamic, and dynamic structures in the solar corona during sunspot minimum and can serve as a diagnostic tool in the interpretation of coronal observations.

Acknowledgments

The contribution of I. Tzur is supported by NASA grant W15-028. The contribution of E. Gliner is supported by grants PHY 81-15800 and PHY 83-13545.

References

- Allen, C. W. 1946, M. N. 106(2), 137.
_____ 1973, Astrophysical Quantities, The Athlone Press.
Baumbach, S. 1937, A. N. 623, 121.
_____ 1939, A. N. 267, 273.
Hoeksema, J. T., J. M. Wilcox, and P. M. Sherrer 1982, JGR. 87 (12).
Hulst, H. C. van de 1950, B.A.N. 11 (410) 135.
Minnaert, M. 1930, Zs. f. Ap. 1, 209.
Moussas, X., N. Papaslamation, V. Rusin, and M. Rybanky 1983, Solar Physics 84, 71-75.
Munro, R., and Jackson, B. 1977, Ap. J. 213, 874.
Osherovich, V.A., Gliner, E. B., Tzur, I., and Kuhn, M. L., 1983, Solar Physics (submitted).
Osherovich, V. A., I. Tzur and E. B. Gliner 1984, Ap. J., 284
Osherovich, V. A. and S. Suess 1982, XXIV COSPAR meeting in Ottawa (abstract).
Saito, K. 1970, Ann Tokyo Astron. Obs. Ser. 2, 12, 53.
Schuster, A. 1880, M. N. 60, 35.
Stenflo, J. O. 1971 Solar magnetic fields, IAU, Symposium 43, 714, D. Reidel, Hingham, Mass: Springer-Verlag, New York.
Uchida, Y., and Low, B. C. 1981, J. Astron. Astrophys. 15, 83.
Withbrow, G. L., T. L. Kohl, H. Weiser and R. H. Munro 1982, Spa. Sci. Rev. 33, 17.
Woltjier, J., Jr. 1926, B. A. N. 3 (94).
Young, R. K. 1911, Lick Obs. Bull. 6 (205) 166.

THEORETICAL APPROACH TO EVOLUTION OF SOLAR MAGNETIC FIELD

Y. Nakagawa

Chiba Institute of Technology, Narashino 275, JAPAN

Abstract. Theoretical approaches to the evolution of solar atmospheric magnetic field are briefly reviewed from the standpoint of their physical significance. A new direction of analysis based on the possible manners of generation of electric current is considered, and its physical implications are discussed.

1. Introduction

The evolution of solar atmospheric magnetic field is clearly governed by the change of photospheric magnetic configuration. This close relationship has been illustrated by a number of analyses, for example, on the basis of evolving force-free magnetic fields by Tanaka and Nakagawa (1973), Nakagawa and Tanaka (1974), Levine and Nakagawa (1974, 1975), and Low and Nakagawa (1975) to name a few. Similar analyses on the basis of evolving magnetostatic fields have been pursued recently by Low (1982a,b; 1984a,b). Apart from such theoretical analyses, pure observational evidences have also been accumulated by Martres, Soru-Escout and Nakagawa (1977) and many others.

The simplest approach to this problem is to examine the evolving potential field for a given photospheric magnetic configuration. Mathematically, this is the boundary value problem, and uniqueness of solution as well as mathematical tools are available for the analysis. However, physically, a current-free potential field is the minimum energy state and to store some magnetic energy in the atmosphere, electric currents must be present. In other words, an evolving potential field cannot be materialized unless the energy gain due to the work done at the photospheric level is exactly dissipated within the atmospheric field.

The success of evolving force-free field analysis suggests the presence of field aligned electric current $J_{||}$ in the solar atmosphere. However, an evolving force-free field can only be achieved by the presence of electric current perpendicular to the field J_{\perp} which is dissipated rapidly to establish the successive force-free fields. The analyses of evolving magnetostatic field accommodate both $J_{||}$ and J_{\perp} ; however, in most studies of this problem, the growth or decay of J_{\perp} is presupposed with little attention to physical change at photospheric level.

In this paper, therefore, this problem is reconsidered from physical grounds, mostly on the basis of generation and dissipation of $J_{||}$ and J_{\perp} .

2. Basic Equations and Their Properties

The basic equations appropriate to the analysis are the magnetohydrodynamic (MHD) equations, i.e., the laws of conservation of mass, momentum and energy, supplemented by the Maxwell equations of the forms:

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{v}) , \quad (1)$$

$$\rho \frac{d\mathbf{v}}{dt} = \mathbf{J} \times \mathbf{B} - \nabla p + \rho \nabla \Phi , \quad (2)$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} , \quad (3)$$

$$\frac{\partial U}{\partial t} = -\nabla \cdot \mathbf{S} , \quad (4)$$

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = \eta \mathbf{J} , \quad (5)$$

and
$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} , \quad (6)$$

where ρ is the density, \mathbf{v} the velocity, \mathbf{J} the current density, \mathbf{B} the magnetic field, p the pressure, Φ the gravitational potential, \mathbf{E} the electric field, U the total energy, \mathbf{S} the energy flux, η the resistivity and μ_0 the permeability, with

$$\frac{d}{dt} = \frac{\partial}{\partial t} + (\mathbf{v} \cdot \nabla) , \quad (7)$$

$$U = \frac{1}{2} \rho v^2 + \frac{1}{\gamma - 1} p + \frac{B^2}{2\mu_0} + \rho \Phi , \quad (8)$$

$$\mathbf{S} = \mathbf{v} \left(U + p + \frac{B^2}{2\mu_0} \right) - \frac{1}{\mu_0} (\mathbf{v} \cdot \mathbf{B}) \mathbf{B} + \frac{\eta}{\mu_0} \mathbf{J} \times \mathbf{B} , \quad (9)$$

and γ being the ratio of specific heats.

It is known (Nakagawa 1981a,b) that the basic set of MHD equations are hyper-

bolic. Hence, the proper physically self-consistent examination of evolving atmospheric magnetic field subject to the (time-dependent) photospheric boundary changes is mathematically an initial-boundary value problem and the analysis must utilize the method of characteristics. Physically, the method of characteristics implies that any change within the medium (including the change on the boundary surface) is related by finite amplitude waves propagating along specific characteristic directions. In other words, a finite-amplitude change on the boundary can be decomposed among these waves and carried to a point in the medium. Consequently, by tracing these waves along the characteristic directions to the source, the causality relationship between the response in the medium and the cause on the boundary can be established uniquely. Nevertheless, practical application of this method is extremely complicated (Sauerwein, 1966); thus a simplified version called the method of projected characteristics has been developed by Nakagawa (1981a,b) and used with success (Han, Wu and Nakagawa, 1982; Wu, *et al.* 1983, 1984).

For the full set of MHD equations, the basic non-linear waves are the fast, slow and transverse (Alfvén) MHD waves, and the entropy wave. The manner of propagation of these waves is not isotropic and depends strongly on the direction of magnetic field as shown in Figure 1, in which the manner of propagation for a point disturbance

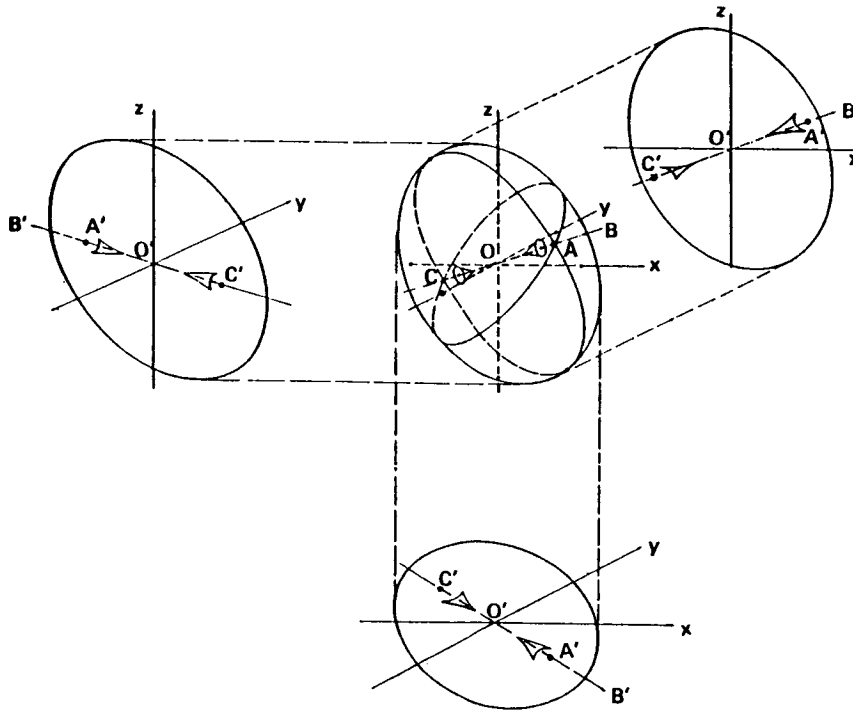


Figure 1. The relationship between three- and two-dimensional loci of characteristic surfaces. The direction of magnetic field in the (x, y, z) -coordinates is represented by OB and the projections are denoted by $O' B'$. A' and C' are the projections of A and C , i.e., the wave fronts of transverse (Alfvén) waves.

at the origin 0 is illustrated for a stationary medium, i.e., for a stationary entropy wave $y = 0$. In Figure 1, OB denotes the direction of the magnetic field and the spheroidal surface (the projected ellipses in the component planes) represents the fast MHD wave front, while a set of conical surfaces along OB axis (projected as cusps) denotes the slow MHD wave fronts and the points A, C along the OB axis (projected as A', C') are the responses by the transverse MHD waves to the point disturbance at 0. It can be seen from this figure that apart from the point to point response by transverse (Alfvén) waves, the deformations of a point source disturbance for fast and slow MHD wave fronts imply that fast and slow MHD waves accompany physical changes due to either expansion or compression. It is important to point out that the transverse MHD (Alfvén) waves only induce a rotation of magnetic field and this physical understanding is verified by Han, Wu and Nakagawa (1982). Thus in a pure two-dimensional plane analysis, contrary to the popular understanding, the transverse (Alfvén) waves are completely dropped out and the magnetic field change is achieved by the fast and slow MHD waves. Under general circumstances, however, a quantitative estimate of a given perturbation among those waves is difficult. Therefore, to gain a physical insight into the nature of evolving magnetic field, a more practical method of approach to this problem is advisable as discussed in the following sections.

3. Physical Approach

Application of a curl-operator to Equation (3) yields, with Equation (6),

$$\begin{aligned}
 \mu_0 \frac{\partial \tilde{J}}{\partial t} &= -\tilde{\nabla} \times \tilde{\nabla} \times \tilde{E} \\
 &= -\tilde{\nabla} (\tilde{\nabla} \cdot \tilde{E}) + \tilde{\nabla}^2 \tilde{E} \\
 &= \tilde{\nabla}^2 \tilde{E} ,
 \end{aligned} \tag{10}$$

where to the MHD approximation, the condition of the prevalence of charge neutrality is used, i.e.,

$$\tilde{\nabla} \cdot \tilde{E} = 0. \tag{11}$$

For the study of solar atmospheric field, we may neglect the effect of resistivity and write

$$\tilde{E} = -(\tilde{v} \times \tilde{B}). \tag{12}$$

Substitution of Equation (12) into Equation (10) gives for the components parallel (\parallel) and perpendicular (\perp) to the magnetic field,

$$\frac{\partial \tilde{J}_{||}}{\partial t} = - \frac{1}{\mu_0} \nabla^2 (\tilde{\mathbf{v}} \times \tilde{\mathbf{B}})_{||} , \quad (13)$$

$$\frac{\partial \tilde{J}_{\perp}}{\partial t} = - \frac{1}{\mu_0} \nabla^2 (\tilde{\mathbf{v}} \times \tilde{\mathbf{B}})_{\perp} . \quad (14)$$

Equations (13) and (14) indicate for a given set of velocity and magnetic fields, whether the evolving magnetic field approaches or deviates from a force-free field.

For an example, let us consider an initial linear potential dipole field subject to a photospheric motion as illustrated in Figure 2a and b. The magnetic field of

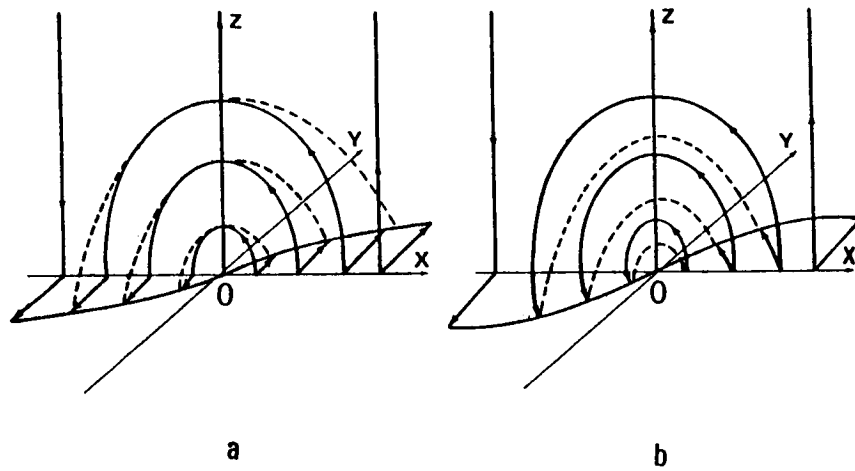


Figure 2. A schematic illustration of physical situation for the evolution of two-dimensional dipole subject to a pure lateral shearing motion (a), and a more general motion (b), of the magnetic foot-points.

the initial linear potential dipole field is confined in the x, z -plane. Then for a pure shearing motion, i.e., for the velocity field only in the y -direction, as in Figure 2a, the immediate response is the x and z -components of $\tilde{\mathbf{v}} \times \tilde{\mathbf{B}}$. In other words, only $\tilde{J}_{||}$ can result; thus, the evolving field must become a force-free field. This expectation is verified by actual computation by Wu *et al.* (1983) in which a pure shearing motion reproduces an evolution close to the analysis by Low and Nakagawa (1975).

Now if the motions are purely convergence or divergence, namely, in the x, z plane, the vector product $\mathbf{v} \times \mathbf{B}$ has only the y -component, hence, only the growth of J_{\perp} results and the evolving field deviates strongly away from a force-free field. It is, therefore, evident that for the photospheric motion shown in Figure 2b, with the growth of both J_{\parallel} and J_{\perp} , the magnetic energy in the atmospheric field increases faster than the case a.

For a complete analysis the detailed response must be worked out at each locality. Nevertheless, the additional condition of prevailing charge neutrality in MHD approximation gives

$$\nabla \cdot \mathbf{J} = 0. \quad (15)$$

This equation can be used to trace the boundary effect into the atmosphere, as the electric currents generated on the boundary must be continuous throughout the medium, i.e., the boundary currents can be used to guide the change of electric current in the upper atmosphere.

The above argument shows that a vortex motion around a circular sunspot tends to produce a force-free field as examined by Nakagawa and Stenflo (1979) and that motions around sunspots favor the development of force-free vortex fields, while intruding or colliding sunspots enhance J_{\perp} . Naturally, the evolution must involve physical changes. Thus let us look at this problem.

A vectorial multiplication of \mathbf{B} to Equation (2) yields

$$\mathbf{J}_{\perp} = \frac{\mathbf{B}}{B^2} \times \left(\rho \frac{d\mathbf{v}}{dt} + \nabla p - \rho \nabla \Phi \right)_{\perp}. \quad (16)$$

For slow change, dropping the term $\rho (d\mathbf{v}/dt)_{\perp}$, we find the magnitude of J_{\perp} to support the prominences or filaments from the deviation of $(\nabla p - \rho \nabla \Phi)_{\perp}$. To the same approximation (i.e., $\rho (d\mathbf{v}/dt)_{\parallel} = 0$), scalar multiplication of \mathbf{B} to Equation (2) gives

$$\nabla_{\parallel} p - \rho \nabla_{\parallel} \Phi = 0, \quad (17)$$

namely, the prevalence of hydrostatic equilibrium along a field line.

In other words, the change in density stratification, such as, formation of a filament or prominence, must involve J_{\perp} . Thus, the next question is in non-static state in which the current J_{\parallel} or J_{\perp} is responsible for the change through $\rho (d\mathbf{v}/dt)_{\parallel}$ or $\rho (d\mathbf{v}/dt)_{\perp}$. For example, restoring the term $\rho (d\mathbf{v}/dt)_{\parallel}$ in Equation (17), we can estimate the magnitude of v_{\parallel} by

$$\tilde{v}_{||} = - \int_0^t \left(\frac{1}{\rho} \tilde{\nabla}_{||} p - \tilde{\nabla}_{||} \Phi \right) dt, \quad (18)$$

4. Discussion and Concluding Remarks

The complex physical changes accompanying an evolution of the atmospheric magnetic field must be treated in a self-consistent manner by the method of characteristics. However, as mentioned in the previous section, a physical approach must be explored, particularly for slow evolution. Naturally, for a fast evolution, such as changes following the onset of a flare or coronal transients the dynamical responses by nonlinear waves must be considered. However, for a slow evolution, it is conceivable that apart from the entropy wave which propagates with the material velocity, the nonlinear waves which propagate rapidly throughout the medium help to achieve the quick adjustment to a quasi-equilibrium state. Then the successive steady state analysis, such as to follow topologically continuous force-free or magnetostatic field can be considered a good approximation to the evolving atmospheric field.

Apart from such considerations the problem of fundamental importance is the prediction of the onset of catastrophic phenomena, flares. From the theory of chaos it is known that for weak nonlinear couplings, an evolving phenomenon reaches a steady final state, while in the presence of a strong nonlinear coupling the final state is unpredictable, i.e., catastrophic. In the solar atmosphere, it is apparent that strong material motions can only be generated by the term $\mathbf{J} \times \mathbf{B}$, that is by the growth of $\tilde{J}_{||}$. At the same time, it is known from the analysis of evolving force-free field that $\tilde{J}_{||}$ cannot exceed a certain limit (Low, 1977) to be stable. Therefore, for the possible onset of a flare, it is important to clarify whether $\tilde{J}_{||}$ or \tilde{J}_{\perp} acts as the trigger for a flare from observational analysis.

For the flare problem, it may be worthwhile to add the plasma physical considerations. In plasma physics, with ρ_e denoting the charge density, Equation (15) becomes the equation of conservation of electric charge

$$\frac{\partial \rho_e}{\partial t} + \tilde{\nabla} \cdot \tilde{\mathbf{J}} = 0, \quad (19)$$

while the equation for electric field is

$$\epsilon_0 \tilde{\nabla} \cdot \tilde{\mathbf{E}} = \rho_e, \quad (20)$$

with ϵ_0 denoting the capacitivity. With Equation (20), Equations (13) and (14) can be rewritten in the forms

$$\frac{\partial \tilde{J}_{||}}{\partial t} = - \frac{1}{\epsilon_0} \nabla_{||} \rho_e - \frac{1}{\mu_0} \nabla^2 (\tilde{\mathbf{v}} \times \tilde{\mathbf{B}})_{||} , \quad (21)$$

$$\frac{\partial \tilde{J}_{\perp}}{\partial t} = - \frac{1}{\epsilon_0} \nabla_{\perp} \rho_e - \frac{1}{\mu_0} \nabla^2 (\tilde{\mathbf{v}} \times \tilde{\mathbf{B}})_{\perp} . \quad (22)$$

In other words, $\tilde{J}_{||}$ as well as \tilde{J}_{\perp} can be generated by the presence of charge density gradient. This supplies an interesting as well as important possibility for the onset of a flare-like instability. All field lines must go through the transition region between the corona and the chromosphere where a considerable variation of electric charge carriers exists. Thus a strong disturbance at these layers can strongly affect the local charge density, such as a formation of a double layer resulting in a large change in the electric current, hence the evolution of the magnetic field.

In summary, up to the present the evolution of magnetic field has been examined mostly from the MHD approximations. Clearly, the MHD approximation is valid for large scale changes; nevertheless, a more systematic analysis must be performed to learn the roles of different physical mechanisms acting in the observed evolution of the magnetic field. It is suggested that examination of the evolution of electric current is a physically promising direction of research and that plasma physical considerations can also provide a new insight to the problem.

References

- Han, S. M., Wu, S. T. and Nakagawa, Y.: 1982, *Computer and Fluids*, 10, 159.
 Levine, R. H. and Nakagawa, Y.: 1974, *Astrophys. J.*, 190, 703.
 Levine, R. H. and Nakagawa, Y.: 1975, *Astrophys. J.*, 196, 859.
 Low, B. C.: 1977, *Astrophys. J.*, 212, 234.
 Low, B. C.: 1982a, *Astrophys. J.*, 254, 796.
 Low, B. C.: 1982b, *Astrophys. J.*, 261, 351.
 Low, B. C.: 1984a, *Astrophys. J.*, (submitted)
 Low, B. C.: 1984b, *Astrophys. J.*, (submitted)
 Low, B. C. and Nakagawa, Y.: 1975, *Astrophys. J.*, 199, 237.
 Martres, M. J., Soru-Escout, I. and Nakagawa, Y.: 1977, *Astron. Astrophys.*, 59, 255.
 Nakagawa, Y.: 1981a, *Astrophys. J.*, 247, 707.
 Nakagawa, Y.: 1981b, *Astrophys. J.*, 247, 719.
 Nakagawa, Y. and Tanaka, K.: 1974, *Astrophys. J.*, 190, 711.
 Nakagawa, Y. and Stenflo, J. O.: 1979, *Astron. Astrophys.*, 72, 67.
 Sauerwein, H.: 1966, *J. Fluid Mech.*, 25, 17.
 Tanaka, K. and Nakagawa, Y.: 1973, *Solar Phys.*, 33, 187.
 Wu, S. T., Hu, Y. Q., Nakagawa, Y. and Tandberg-Hanssen, E.: 1983, *Astrophys. J.*, 266, 866.
 Wu, S. T., Hu, Y. Q., Nakagawa, Y. and Tandberg-Hanssen, E.: 1984, *Astrophys. J.*, (submitted).

Current Evolution in a Numerical Emerging-Magnetic-Flux Model

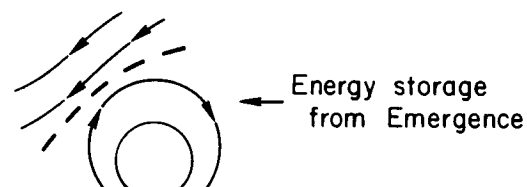
T.G. Forbes

Abstract

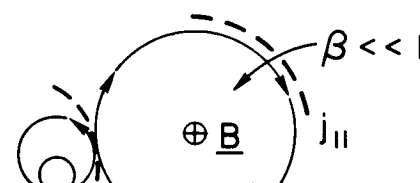
The resistive-MHD equations are numerically solved in two-dimensions for an initial-boundary-value problem which models the emergence of magnetic flux from the photosphere into the corona. As the emergence begins a current sheet forms around the emerging region which separates the emerging region from the overlying coronal magnetic field. This current sheet is the source of the free-magnetic energy in the system, and in the limit of zero resistivity it is a simple tangential discontinuity. However, when the resistivity is finite, reconnection between the magnetic field in the emerging region and the overlying coronal magnetic field ensues, and the subsequent evolution of the enveloping current sheet becomes complex. The overall time history of the current evolution is suggestive of the expected current evolution for the pre-flare, impulsive, and main phases of flares.

1. Introduction

1) Emergence of Potential Field

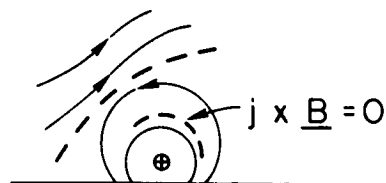


2) Emerging Flux as Trigger



3) Combinations

a. Force-free and Current Sheet



b. Non-force-free and Current Sheet

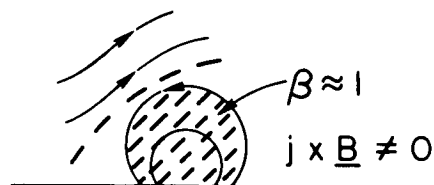


Figure 1. Some examples of possible flare mechanisms involving emerging magnetic flux. Case 1-simple current sheets model, Case 2-emerging flux as flare trigger, and Case 3-emerging flux with force-free (a) or non-force-free (b) internal currents.

Many solar flares are seen to occur after the emergence of new magnetic flux from beneath the photosphere (Martres et al., 1968, Rust, 1972; Vorpahl, 1973; Sheely et al., 1975; Glackin, 1975; Martin and Harvey, 1973), and various analytical models have been developed to explain this occurrence (Canfield et al., 1974; Heyvaerts et al., 1977; Tur and Priest, 1976, 1978; Milne and Priest, 1981). One aspect that is particularly difficult to treat analytically is the reconnection between the field of the emerging flux region and the pre-existing field in the corona. One way to avoid the analytical difficulties is to solve the magnetohydrodynamic equations numerically, but there is a serious handicap in such an approach, namely the inability of present day computers and techniques to handle simultaneously the very large and very small physical scales existing in the solar atmosphere. Nonetheless, numerical solutions are still very useful as a tool with which to explore the nonlinear dynamics implied by the governing equations.

There are several scenarios one might consider for an emerging flux trigger for a flare. Examples of some simple possibilities are shown in Figure 1. Case No. 1 shows the simplest possible case of the emergence of a potential field in which all the free-magnetic energy available for the flare is stored in the current sheet. However, one might consider the alternate possibility shown in case 2 where the flare magnetic energy is already stored in pre-existing, force-free structures in the corona. In this case the emergence would not itself supply any significant energy, but would only serve to trigger an instability in the pre-existing structure. Combinations of the type shown in case 3 are also possible. In these scenarios the emerging flux region may itself contain either force-free currents (a) or distributed currents (b) depending upon the plasma beta of the emerging region.

For the present two-dimensional calculation the scenario shown in 3b has been used as a basis for the following initial-boundary-value problem:

2. Mathematical Formulation

2.1. Governing equations

The partial differential equations for the problem are:

$$\partial \underline{B} / \partial t = \nabla \times (\underline{V} \times \underline{B}) + \eta \nabla^2 \underline{B}. \quad (1)$$

$$\partial \rho / \partial t = -\nabla \cdot (\rho \underline{V}). \quad (2)$$

$$\rho [\partial \underline{V} / \partial t + (\underline{V} \cdot \nabla) \underline{V}] = -\nabla p + \underline{j} \times \underline{B} \quad (3)$$

$$\partial p / \partial t + (\underline{V} \cdot \nabla) p = -\gamma p \nabla \cdot \underline{V} + (\gamma - 1) \eta j^2 \quad (4)$$

$$\underline{j} = \nabla \times \underline{B} \quad (5)$$

$$p = \rho T \quad (6)$$

\underline{B} , \underline{V} , ρ , p , \underline{j} , \underline{x} , t are dimensionless variables defined by

$$\begin{aligned}
\underline{B} &= \underline{B}'/B'_0, & \rho &= \rho'/\rho'_0, & \underline{V} &= \underline{V}'/V'_a, \\
p &= 4\pi p'/B'_0{}^2, & \underline{j} &= \underline{j}'4\pi L'_0/(B'_0 c), & V'_a &= B'_0/(4\pi\rho'_0)^{1/2}, \\
\underline{X} &= \underline{X}'/L'_0, & t &= t'V'_a/L'_0,
\end{aligned}$$

where c is the speed of light and \underline{B}' , \underline{V}' , ρ' , p' , \underline{j}' , \underline{X}' , and t' are the dimensional magnetic field, flow velocity, mass density, pressure, current density, spatial coordinate vector, and time, respectively. The scaling constants B'_0 and ρ'_0 are defined so as to normalize the dimensionless variables with respect to the initial conditions, and L'_0 is the size of the computational domain. γ is the ratio of specific heats ($=5/3$), and η is the dimensionless magnetic diffusivity ($=0.0005$). η is related to the dimensional magnetic diffusivity, η' , by

$$\eta = \eta'/(L'_0 V'_a),$$

which is the inverse of the magnetic Reynolds number (or, more properly, the inverse of the Lundquist number).

2.2 Initial conditions

The initial conditions are:

$$\underline{B} = \hat{e}_x, \quad (7)$$

$$\underline{V} = \underline{0}, \quad (8)$$

$$\rho = 1, \quad (9)$$

$$p = \beta/2, \quad (10)$$

where β , the initial ratio of gas to magnetic pressure, is 0.1.

2.3 Boundary conditions

Formulation of the boundary conditions presents a considerably more complex task than that of choosing the initial conditions. The numerical calculation is executed over a uniform cartesian grid mesh of 49×97 points enclosed by four boundaries. Of these four boundaries only the one corresponding to the base of the corona ($Z = 0$) is physically real. The other three boundaries are non-physical and are imposed by the finite spatial domain of the numerical grid. In our coordinate system these three non-physical boundaries correspond to the $X = 0.0$ plane (left side), the $Z = 1.0$ plane (top), and the $X = 1.0$ plane (right side) of the square computational domain. To reduce the number of mesh points needed to achieve a given resolution, the $X = 0$ boundary (left side) is made an axis of symmetry. For the two remaining non-physical boundaries we use the simple extrapolation method discussed by Orlanski (1976) in order to simulate an open boundary through which magnetic flux and plasma can enter or exit from the box.

The physically important boundary conditions for the emerging flux process are the ones at the base of the computational domain. The nature and number of these conditions depend upon whether one considers the reduced hyperbolic set of equations obtained by neglecting all dissipation terms, or whether one considers the general parabolic set which includes dissipation terms (Chu, 1978). Here we consider the parabolic system obtained when magnetic diffusivity is added to the ideal-MHD system. For such a system the number of independent boundary conditions that are needed depends upon the number of characteristics propagating from the boundary into the fluid, and this in turn depends upon the speed of the flow through the boundary (e.g. Steinolfson and Nakagawa, 1976).

Virtually all explicit numerical methods require each variable to be specified at the boundary. If the number of physically allowed independent boundary conditions is less than the total number of variables, then additional 'compatibility' conditions are needed (Nakagawa, 1981). In general these compatibility conditions depend both upon the other boundary conditions and the initial conditions, and only in some especially simple cases can they be analytically determined. Otherwise, they must be numerically computed at each time step (for example, see Han et al., 1982).

For our initial-boundary-value problem it is convenient to consider the boundary conditions inside and outside the emergence region separately. We assume that the edge of the emerging region is defined by a given time-dependent function, $X_e(t)$, and for $X > X_e(t)$ the independent boundary conditions are:

$$B_z = 0. \quad (11a)$$

$$\rho V_z = 0. \quad (12a)$$

$$\partial B_x / \partial Z = 0. \quad (13a)$$

with the dependent boundary conditions (i.e. compatibility relations):

$$\partial (\rho V_x) / \partial Z = 0. \quad (14a)$$

$$\partial \rho / \partial Z = 0. \quad (15a)$$

$$\partial p / \partial Z = 0. \quad (16a)$$

These are the conditions for a grounded wall with no magnetic flux threaded through it.

For the boundary conditions inside the emerging region we consider the times before and after emergence individually, starting with the more easily formulated time-independent ones occurring after the emergence is complete. Defining t_e , as the time when flux stops emerging, we have for $t > t_e$ and $X < X_e$ the independent boundary conditions:

$$B_z = f(X). \quad (11b)$$

$$\rho V_z = 0. \quad (12b)$$

$$\partial B_x / \partial Z = \partial B_z / \partial X. \quad (13b)$$

$$\rho V_x = 0. \quad (14b)$$

with the dependent boundary conditions:

$$p/\rho^\gamma = k_1 = \text{constant}. \quad (15b)$$

$$\partial p / \partial z = 0. \quad (16b)$$

Here, instead of a compatibility relation for ρV_x , we have used an independent, no-slip boundary condition (14b). This extra boundary condition is mathematically permissible if viscosity is incorporated into the MHD equations. Since only a numerical viscosity is present in our system of equations, this could give rise to a poorly resolved viscous boundary layer were a substantial tangential flow to appear at the boundary in this region. However, owing to the very small value of the electrical resistivity, η , only an insubstantial diffusive tangential flow component can occur in the emergence region, and hence no viscous boundary layer appears. The use of the no slip condition greatly simplifies the formulation of the boundary conditions which in turn decreases the cost of the calculation significantly. (Otherwise, additional numerical iterations would be required to solve the compatibility relations at each timestep.) The condition (13b) together with (14b) serve to anchor or 'line-tie' the magnetic field to the base so that field lines cannot move along this boundary.

Inside the emerging region ($X < X_e$) and during the emergence ($t < t_e$) the independent boundary conditions are:

$$B_z = f(X, t). \quad (11c)$$

$$\rho V_z = g(X, t). \quad (12c)$$

$$\partial B_x / \partial Z = \partial B_z / \partial X. \quad (13c)$$

$$\rho V_x = 0. \quad (14c)$$

$$p/\rho^\gamma = k_1 \quad (15c)$$

with the compatibility relation:

$$\partial \rho / \partial t + \partial (V_z \rho) / \partial Z = 0. \quad (16c)$$

Two additional independent boundary conditions (14c and 15c) are required when an inflow occurs since the entropy (15c) and tangential momentum (14c) of the incoming fluid must be specified (Chu, 1978). However, there is no simple expression for density (or pressure) in this case, and so a time-dependent

compatibility relation (16c) is now needed. To solve (16c) we use the one-sided differencing scheme described in Chu and Sereny (1974).

The condition (13c) is based on the assumption that there is no diffusion of the magnetic field through the incoming fluid prior to its entry into the numerical box. In other words we assume, as we do after the emergence, that the frozen-flux conditions, $\underline{E} = -\underline{V} \times \underline{B}$, holds beneath the base. This is consistent with the assumption of absolute line-tying of the field to the fluid, in the region below the corona (i.e. below the transition zone) at all times and locations. In reality we expect this assumption only to be approximately true since the photospheric line-tying is inertial ($\underline{V} \times \underline{B} = 0$) and not both inertial and resistive ($\underline{V} \times \underline{B} + \eta \mathbf{j} = 0$). However, to incorporate such partial line-tying would require a model of the transition zone, and this is beyond the scope of the present simulation.

To complete the description of the boundary conditions along the base we specify the functional forms of $f(X,t)$ and $g(X,t)$ along with the associated constants. Our choice of these functions and constants is based on the concept of a shielded magnetic flux tube moving upward through the base. The field inside the flux tube is chosen to be a simple sinusoid, namely:

$$f(X,t) = \begin{cases} B_e X(X^2+h^2)^{-1/2} \sin[\pi(X^2+h^2)^{1/2}/W], & 0 \leq t < t_e \\ B_e \sin(\pi X/W), & t \geq t_e. \end{cases} \quad (17)$$

The upward vertical momentum is chosen on the assumption that beneath the base the flux tube is moving upwards at a known speed $\dot{h}(t)$, and that it has a density distribution obtained by solving the pressure balance equation

$$\partial p / \partial r + \partial (B_\theta^2/2) / \partial r + B_\theta^2/r = 0, \quad (18)$$

where r is the radial distance inside the tube and $B_\theta(r)$ is the poloidal magnetic field. It should be emphasized that (18) is not itself used as a boundary condition but only as an aid in choosing the inflow momentum function $g(X,t)$. With (18) we obtain

$$g(X,t) = \beta^{-1} \dot{h} B_e^2 [C_1 (2\pi[X^2+h^2]^{1/2}/W) - \ln(\pi[X^2+h^2]^{1/2}/W) - \sin^2(\pi[X^2+h^2]^{1/2}/W)] + 2P^*, \quad (19)$$

where

$$h(t) = \begin{cases} -W \cos[\pi t/(2t_e)], & 0 \leq t < t_e, \\ 0, & t \geq t_e. \end{cases}$$

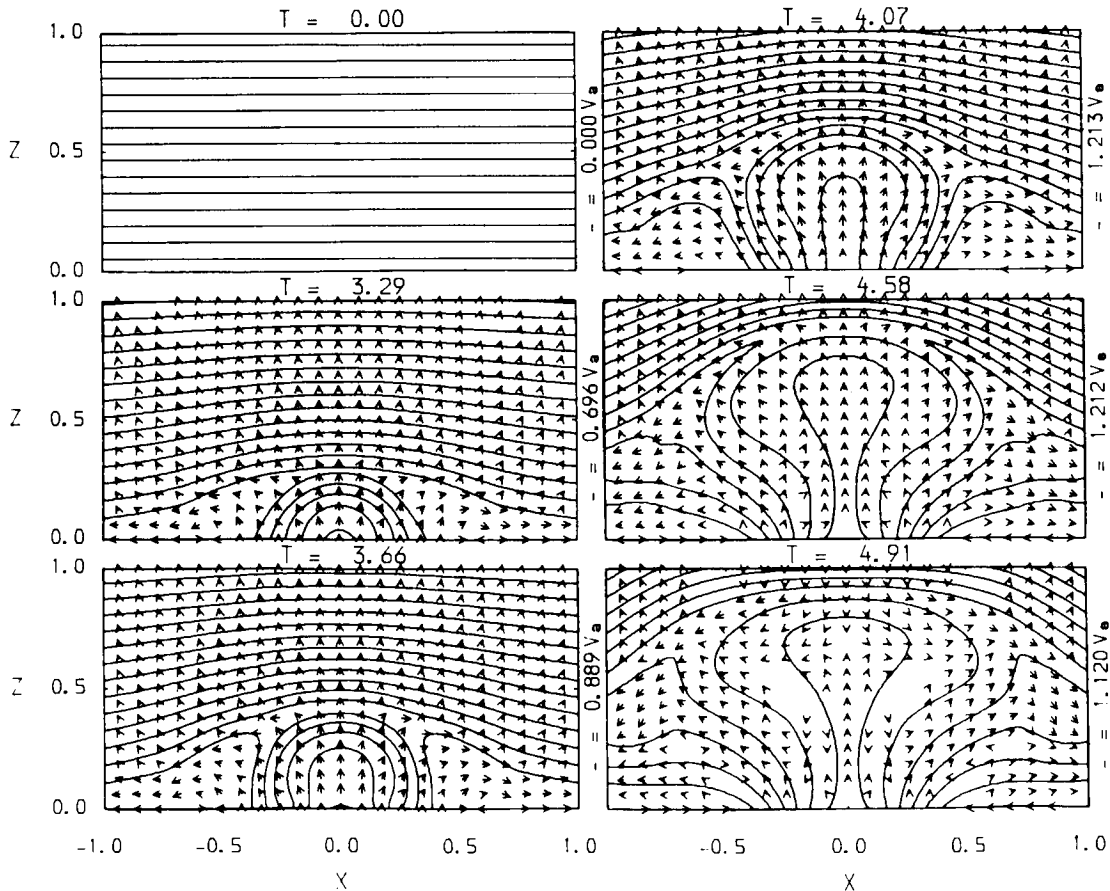
The intersection of the flux tube with the base defines the edge of the emerging flux region, i.e.

$$X_e = [W^2 - h(t)^2]^{1/2}.$$

The various constants in the above expressions are set as

$$\begin{aligned} k_1 &= \beta/2, \quad P^* = 1.60, \quad B_e = 1.34, \\ \beta &= 0.1, \quad t_e = 4.0, \quad W = 0.5, \\ \gamma &= 5/3. \end{aligned}$$

The above constants have been chosen on the basis that beneath the base the tube has a uniform entropy distribution equal to the initial entropy in the box. The half-width of the tube is W , the maximum field strength inside the tube is B_e , and β is the external plasma beta parameter. $P + 0.59 B^2$ is the pressure at the surface of the tube when it is beneath the base, and it is set equal to the initial total pressure (gas + magnetic) in the numerical box. (It is not necessarily the pressure at the surface of the tube when the tube is at or above the base, since then the pressure is determined by the numerical solution.) h gives the location of the axis of the tube below the base, and t_e is the time required for the centre of the tube to reach the base. After t_e , the flow of plasma and magnetic flux through the base is stopped. It is found in practice that the inflow velocity is nearly uniform along the base and corresponds to an injection speed, h , of $\sim 1/8$ times the scale (i.e. coronal) Alfvén velocity.



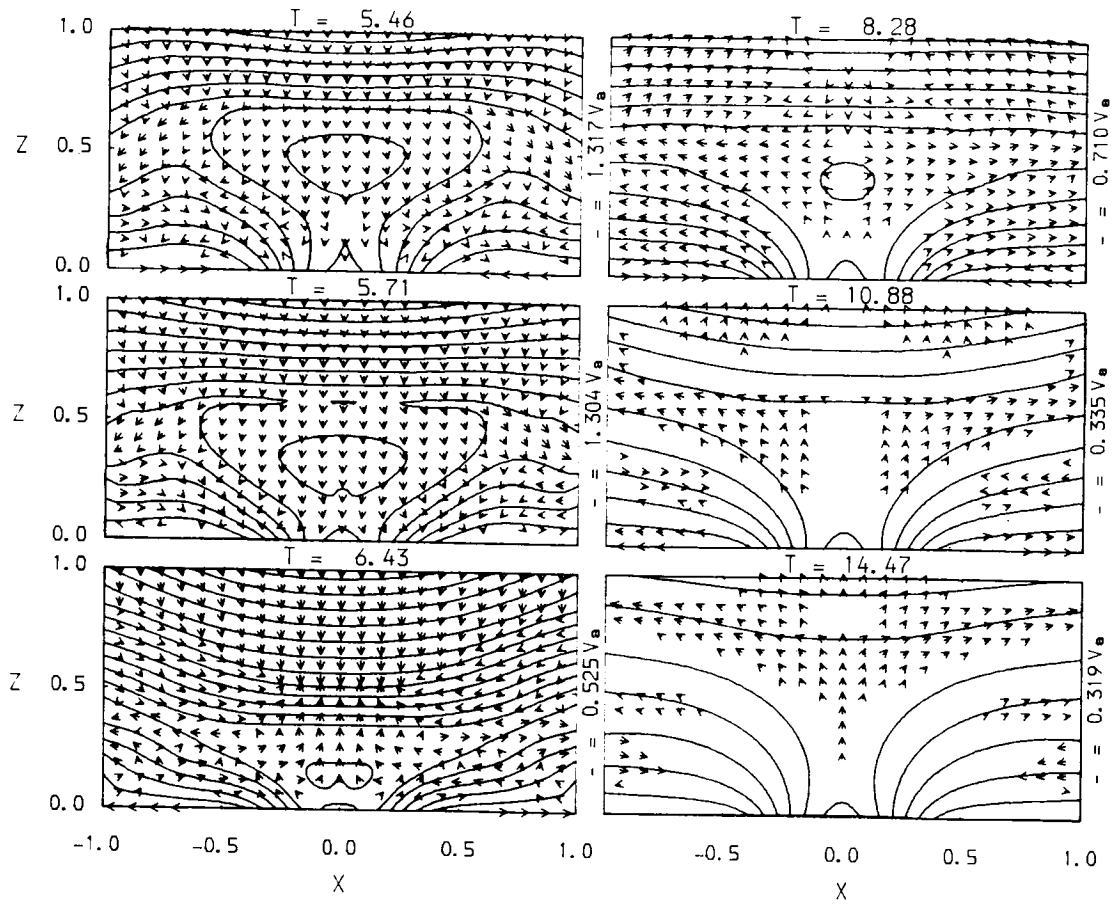


Figure 2. Time sequence of magnetic field lines and flow vectors showing reconnection between a pre-existing, horizontal magnetic field and an emerging region of closed magnetic field lines which are antiparallel to the pre-existing ones. The times are in units of the time initially required for an Alfvén wave to travel from the bottom to the top of the numerical box. The magnitudes of the flow vectors are in units of the initial Alfvén velocity, V_a , and are rescaled in each panel with the scaling of the fastest flow speed indicated to the right of each panel. Emergence of new flux at the base occurs in the region $|x| < 0.5$ from $t = 0.0$ to $t = 4.0$ at an input speed of about $0.12 V_a$.

The boundary conditions set the entropy at the base equal to its initial value, and thus, the temperature at the base is

$$T = 0.5 \beta \rho^{2/3}.$$

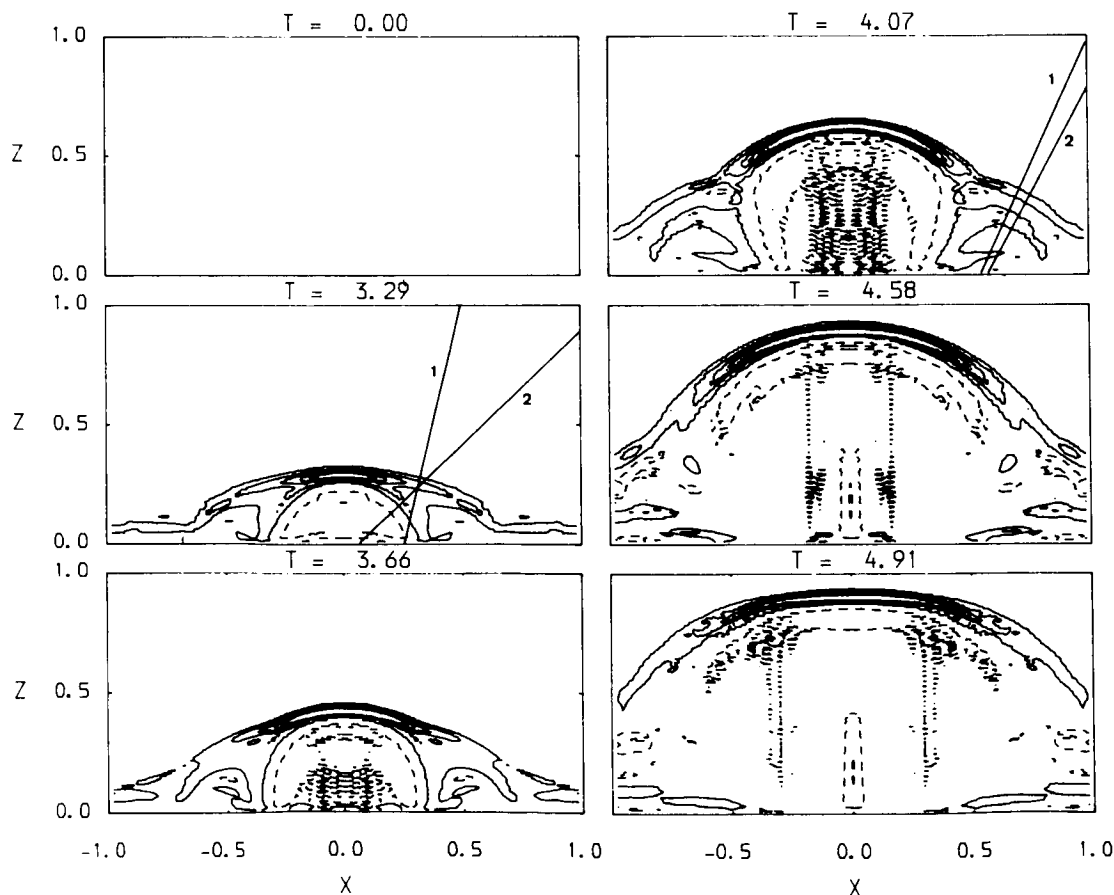
The temperature of the fluid entering through the base is about twice the initial temperature in the box.

2.4 Numerical Method

Equations (1) - (5) together with the initial conditions (7) - (10) and the boundary conditions (11) - (16) are solved by a method based upon the explicit, flux-corrected transport (FCT) code SHASTA (SHarp And Smooth Transport Algorithm) developed by Boris and Book (1976). Our version of the code is similar to Weber et al. (1979) except that we use the magnetic vector potential rather than the magnetic field components in that section of the code which diffuses the field. For the 49×97 grid point mesh used in our calculation the effective numerical viscous and magnetic Reynolds numbers are on the order of 10^6 outside discontinuities (e.g. shocks), but to stabilize the code against numerical instabilities triggered by the growth of small scale-length oscillations it is necessary to include an explicit magnetic diffusion. This diffusion required for stabilization corresponds to a magnetic Reynolds number of about 10^3 (see Matthaeus and Montgomery, 1981). The calculation is carried out for approximately 15 Alfvén scale times and requires about 50 hours of cpu time on a VAX-11/780.

3. Numerical Results

Figure 2 shows the evolution of the magnetic field and flow for the initial-value-boundary problem set forth above. By $t = 3.29$ (middle left panel) the emergence is almost complete, and only a relatively small amount of



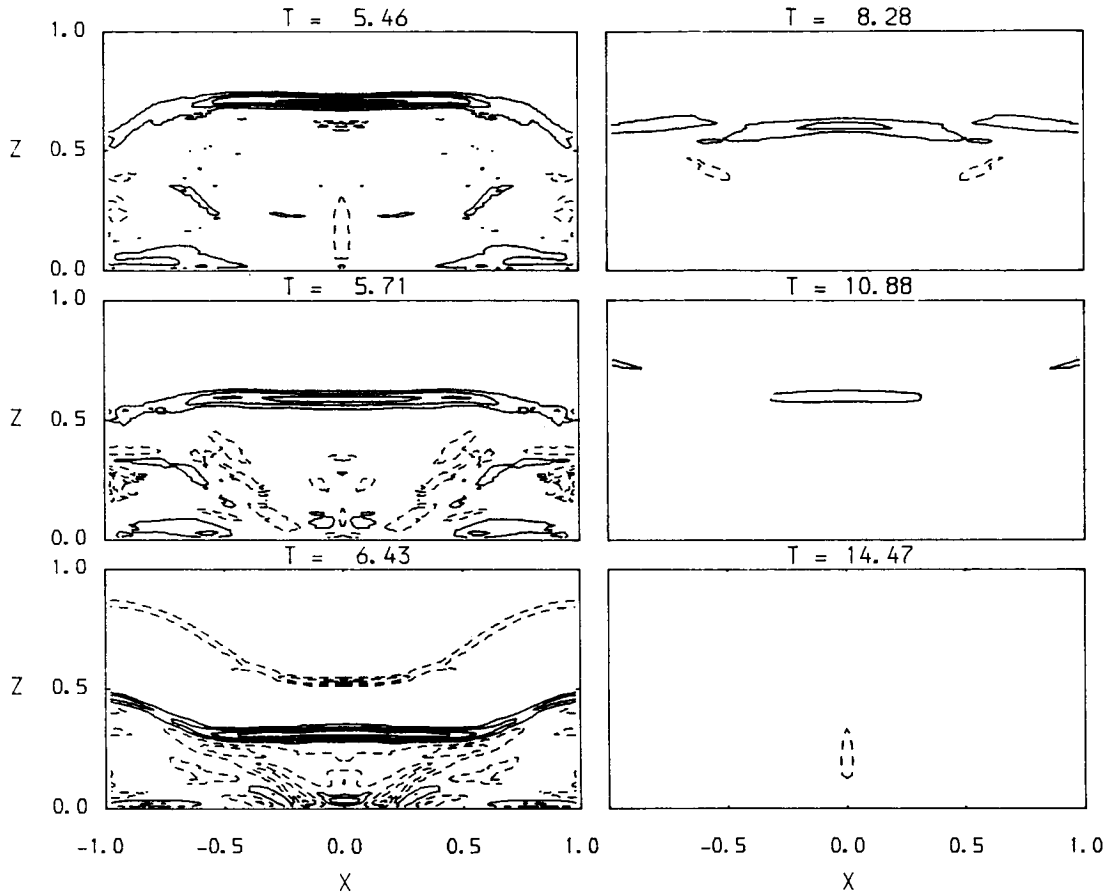


Figure 3. Current density contours for the same times shown in Figure 2. Solid lines indicate regions with positive current density, and dashed lines indicate regions with negative current density. The units are normalized with respect to $B_0/(2W)$, the initial magnetic field over the width of the emerging region, and the jump, Δj , between adjacent contour lines is 13.3. The straight lines labelled 1 and 2 in the panels $t = 3.29$ and 4.07 designate normals to the upper (1) and lower (2) current sheets through which they pass. The choppy, chaotic contour lines in the centre of the emerging flux region during $t = 3.66$ to 4.91 are due to the development of fluctuations at scale lengths smaller than the grid spacing.

reconnection has taken place. At this time the strongest flows occur in the region between the emerging flux field and the initial overlying field (*i.e.* in the reconnected region), but by $t = 3.66$ the strongest flows occur in the middle of the emerging flux region (*i.e.* in the unreconnected region). The upward flow in the middle of the emerging flux region has a speed of ~ 0.9 which is considerably greater than the injection speed of ~ 0.12 . This high speed flow is a result of an imbalance that develops between magnetic ($\mathbf{j} \times \mathbf{B}$) and pressure (∇p) forces inside the emerging flux region. The flow causes the emerging flux region to balloon upwards and outwards even after the injection of flux is completed (*c.f.* $t = 4.07 - 4.91$), and as a result of the ballooning a magnetic

island is formed as the stretched field lines near the base begin to reconnect (c.f. $t = 5.46 - 8.28$).

The upward rush of fluid ($t = 5.46$) is resisted by the overlying magnetic field, and by $t = 4.91$ the resulting tension in the overlying field drives a strong downflow that creates a fast-mode MHD shock ($t = 6.43$ along $Z \sim 0.5$). After the fast shock propagates out the top, the system settles down to a quasi-steady configuration containing, a more or less, stationary magnetic island. Due to continuing reconnection this island eventually disappears ($t \sim 14.0$), and the configuration becomes virtually potential ($\nabla \times \underline{B} = 0$) with little kinetic energy remaining in the computational domain.

The shocks generated during the evolution of the system are most apparent in the current density contour plots shown in Figure 3. At all times the strongest current density occurs at the locations of the x-type neutral lines (x-lines). If there were no reconnection at all, we would expect the current density in the vicinity of the first x-line (panels $t = 3.29$ to 4.07) to form a current sheet of the tangential discontinuity type discussed by Tur and Priest (1976). This current sheet would completely enclose the emerging flux region, separating it from the overlying, horizontal magnetic field. But since reconnection is occurring, what we see in Figure 3 looks considerably more complicated.

Emanating outward from the region near the x-line are two sets of current sheets which enclose the low magnetic field region of newly reconnected field lines (c.f. $t = 3.29$). These are slow-MHD shocks which before $t \sim 3.5$ are roughly similar to those predicted by the standard Petschek picture of reconnection. However, after $t \sim 3.5$, the shock structure changes rapidly as the developing non-equilibrium slams the emerging magnetic flux region into the overlying field. The upper set of slow shocks are left behind, and a new set of slow shocks appear above the older set ($t = 3.66$ to 4.07). By $t = 4.58$ the older set has faded, and again there are just two sets of shocks surrounding the reconnected field regions. (The lower set is difficult to see in the current density contours in Figure 3, but can be seen clearly in Figure 2 by examination of the flow vectors at $t = 4.58$.)

Onwards of about $t = 5.5$ little remains of the earlier reconnection, slow-shock system. After $t = 5.5$, the dominant current density feature is the current sheet separating the regions of emerged and pre-existing magnetic field regions, and only a slight enhancement of the current density at the x-lines is apparent. Also predominant are structures associated with the global field oscillations, as for example the fast-mode shock at $t = 6.43$ (along $Z \sim 0.7$).

One question that seems appropriate to consider before proceeding further is why does the large-scale imbalance between $\underline{j} \times \underline{B}$ and ∇p forces at about $t = 3.6$ occur? It seems to us that there are basically 3 likely possibilities as follows:

1. The emerging region is either not in equilibrium, or it is immediately in an ideal-MHD unstable equilibrium when it enters the box.

2. The emerging region is initially in a stable equilibrium, but either equilibrium or stability is lost simply by the process of continued emergence.
3. The emerging region is in a stable mechanical equilibrium, but it is driven either into non-equilibrium or ideal-MHD instability by reconnection.

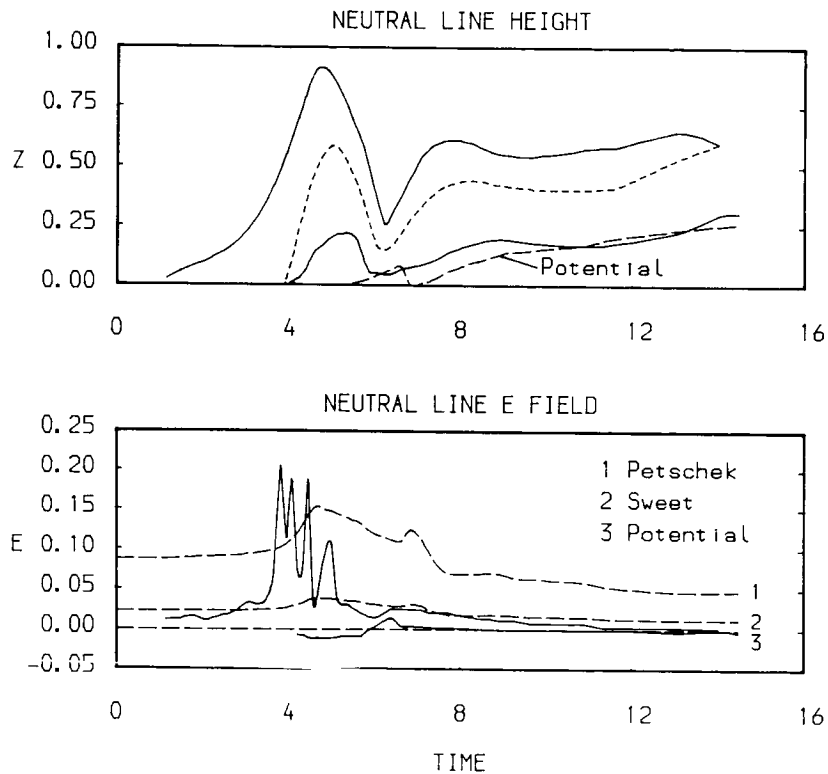


Figure 4. The heights of the x-type neutral lines (solid curves in the upper panel), and their reconnection rates (solid curves in the lower panel) as functions of time. In the upper panel the o-type neutral line at the centre of the magnetic island is indicated by the unlabelled dashed line, and the x-line for the potential solution is indicated by the appropriately labelled dashed line. In the lower panel the dashed lines labelled 1 and 2 indicate the maximum rates of reconnection predicted by the steady-state theories of Petschek (1) and (2), while the dashed line labelled 3 is just $E_o = 0$ which is the rate for the potential solution (i.e. no reconnection).

In Figure 4 neutral line heights (both x- and o-types) and the reconnection rates at the two x-lines are shown as functions of time. Also shown are the height of the x-line obtained in the potential solution (see Appendix B) and the maximum reconnection rates predicted by the steady-state theories of Petschek (1964) and Sweet (1958). For the Sweet theory the electric field at the x-line is

$$E_o = V_a R_m^{-1/2},$$

where V_a is the Alfvén speed in the inflow region at a distance far from the x-line.^a Here we evaluate V_a at $X = 0.0$, $Z = 1.0$, and since this value varies with time, the Sweet value of E_o also varies with time. For the Petschek theory we use the compressible version of Soward and Priest (1982) which gives

$$E_o = V_a [\pi\gamma/(4\gamma-2)] / [\ln(R_m^2 E_o/V_a)] .$$

The time history of the electric field at the neutral line suggests that four phases occur during the evolution of the system. We identify these as a quasi-steady emergence phase from $t = 0.0$ to about $t = 3.5$, an impulsive phase from $t = 3.5$ to 6.0 , a second quasi-steady phase from $t = 6.0$ to 12.0 , and finally, a static, potential phase from $t = 12.0$ onward.

The sudden increase in E_o at the onset of the second phase after the relatively quiescent first phase suggests that the erosion by reconnection of the outer layers of the emerging-flux region eventually upsets the mechanical balance existing between the $\underline{j} \times \underline{B}$ and ∇p forces inside the emerging flux region. However, it is also possible that the emerging region simply becomes ideally unstable once a sufficient quantity of flux has emerged, and so we cannot be certain that reconnection is the cause of the impulsive phase. However, whatever the cause, the fluid is rapidly accelerated upwards and reconnection of the anti-parallel magnetic field is driven at a very fast rate. After a few oscillations a new mechanical equilibrium is achieved, and the rate of reconnection decreases as the flow driving it disappears. Reconnection continues at a reduced level until eventually ($t > 12$) the lowest energy (*i.e.* potential) state is achieved.

4. Aspects Relevant to Solar Vector Magnetic Fields

In the numerical model of the emerging flux region discussed here, the normal (*i.e.* line of sight) component of the magnetic field, B_n , only varies during the process of emergence. Once the emergence is complete, B_n is constant even though large current changes continue to occur. The reason for this is the assumption that field lines are line-tied (*i.e.* anchored) at their photospheric footpoints. Thus, in our model only the transverse components are a function of the currents in the corona.

Because the transverse components of the field are produced only by transverse currents ($\underline{j} \cdot \underline{B} = 0$), our model has no current parallel to the

magnetic field. Unlike many flare models, the parallel currents are not essential to the emerging flux model (see Priest, 1982). In the present numerical model the magnetic energy that drives the flare is stored in the current sheet separating the emerging region from the overlying corona and in the distributed transverse current within the emerging region. (If the plasma beta parameter within the emerging region is small (i.e. magnetic pressure dominant over gas pressure), then the distributed transverse current should be replaced by a force-free current.)

The important point is that the emerging-flux model does not require a current parallel to the field. If observations of the transverse magnetic field components continue to indicate that parallel currents do not decay during some flares (Krall *et al.*, 1982), then the implication would be that the magnetic energy liberated from these flares is derived from transverse current (e.g. current sheets or filaments) such as occur in the numerical simulation presented here.

Acknowledgement

The author is most grateful to the British Science and Engineering Research Council and to the NASA Solar-Terrestrial Program (NASA grant NAGW-76 to the University of New Hampshire) for financial support.

References

- Boris, J.P. and Book, D.L.: 1976, J. Comp. Phys., **20**, 397.
 Canfield, R.C., Priest, E.R. and Rust, D.M.: 1974, in Y. Nakagawa and D.M. Rust (eds.), Flare-related Magnetic Field Dynamics, NCAR, Boulder, CO, U.S.A.
 Chu, C.K.: 1978, Advances in Applied Mech. **18**, 285.
 Chu, C.K. and Sereny, A.: 1974, J. Comp. Phys., **15**, 476.
 Forbes, T.G. and Priest, E.R.: 1982, Solar Phys., **81**, 303.
 Forbes T.G. and Priest, E.R.: 1983, Solar Phys., **84**, 170.
 Glackin, D.L.: 1975, Solar Phys., **43**, 317.
 Han, S.M., Wu, S.T. and Nakagawa, Y.: 1982, Computer and Fluids, **10**, 195.
 Heyvaerts, J., Priest, E.R. and Rust, D.M.: 1977, Astrophys. J., **216**, 123.
 Krall, K.R., Smith, J.B. Jr., Hagyard, M.J., West, E.A., and Cummings, N.P.: 1982, Solar Phys., **79**, 59.
 Martin, S. and Harvey, K.: 1973, Solar Phys., **32**, 389.
 Martres, M.J., Michard, R. Soru-Iscovisi, I. and Tsap, T.: 1968, in K.O. Kiepenheuer (ed.), Structure and Development of Solar Active Regions, IAU Symp., **35**, 318.
 Matthaeus, W.H. and Montgomery, P.: 1981, J. Plasma Phys., **25**, 11.
 Milne, A.M. and Priest, E.R.: 1981, Solar Phys., **73**, 157.
 Nakagawa, Y.: 1981, Astrophys. J., **247**, 719.
 Orlanski, I.: 1976, J. Comp. Phys., **21**, 251.
 Petschek, H.E.: 1964, AAS-NASA Symp. on Solar Flares, NASA SP-50, p.344.
 Priest, E.R.: 1982, Solar Magnetohydrodynamics, Reidel, London.

Rust, D.M.: 1972, Solar Phys., 25, 141.
Sheeley, N.R. and Harvey, J.W.: 1975, Solar Phys., 45, 275.
Soward, A.M. and Priest, E.R.: 1982, J. Plasma Phys., 28, 335.
Steinolfson, R.S. and Nakagawa, Y.: 1976, Astrophys. J., 207, 300.
Sweet, P.A.: 1958, Nuovo Cimento Suppl. 8X, 188.
Tur, T.J. and Priest, E.R.: 1976, Solar Phys., 48, 89.
Tur, T.J. and Priest, E.R.: 1978, Solar Phys., 58, 181.
Vorpal, J.A.: 1973, Solar Phys., 28, 115.
Weber, W.J., Boris, J.P. and Gardner, J.J.: 1979, Computer Phys. Commun.
16, 243.

THE GLOBAL STRUCTURE OF MAGNETIC FIELDS WHICH SUPPORT
QUIESCENT PROMINENCES

Ulrich Anzer
Max-Planck-Institut für Physik und Astrophysik
Institut für Astrophysik
Karl-Schwarzschild-Str. 1
8046 Garching, FRG

Magnetic fields in quiescent (and other) prominences have been observed for a long time. But only recently has it become possible to measure the full magnetic field vector (Sahal-Brechot et al. 1977). The component of the field along the line of sight, $B_{||}$, can be uniquely determined, whereas for the component perpendicular to the line of sight \vec{B}_{\perp} and $-\vec{B}_{\perp}$ are indistinguishable. Therefore there remains an ambiguity in the actual magnetic field vector, and in particular with respect to its orientation relative to the prominence axis. Leroy et al. (1983) have studied a sample of more than hundred prominences. Also Athay et al. (1983) presented a more detailed analysis of 10 prominences. One can then set these prominence fields into relation to the underlying photospheric fields. Leroy (1984) found from his statistical analysis of several hundred prominences that in 25% of the cases the field penetrates the prominence directly, whereas in 75% the field orientation in the prominence is reversed. These results are very crucial for the modelling of prominences.

Basically two different types of models have been developed; one is due to Kippenhahn and Schlüter (1957), the other to Kuperus and Raadu (1974). The Kippenhahn-Schlüter configuration is shown in Fig. 1a. In such a configuration the support of

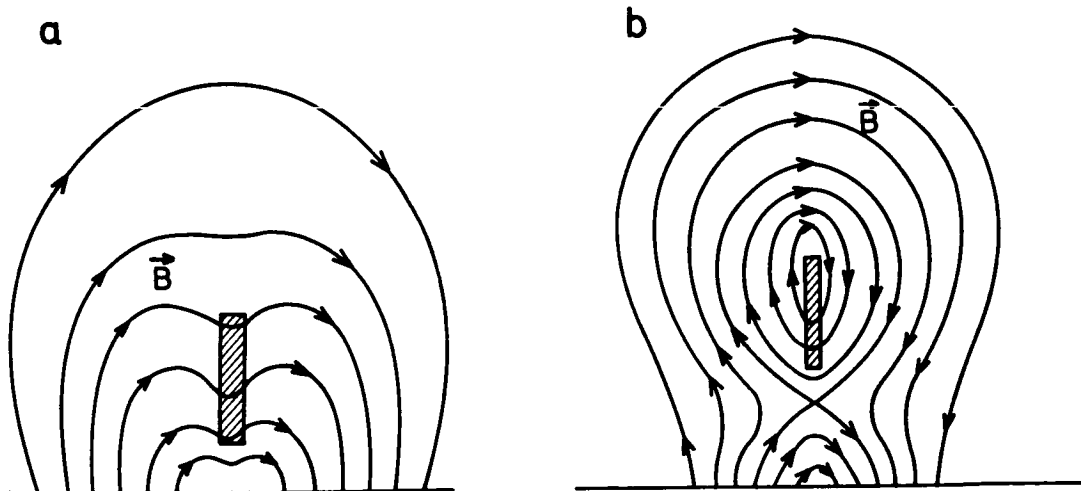


Fig. 1: a) The Kippenhahn-Schlüter configuration
b) The Kuperus-Raadu configuration.

the heavy prominence material presents no problem, although some questions about the stability and the thermal insulation are still not completely solved. In the configuration shown the electrical current associated with the prominence flows out of the plane drawn in Fig. 1a. In the Kuperus-Raadu configuration the orientation of the horizontal field component through the prominence is reversed (Fig. 1b). In such a model the field lines penetrating the prominence are all closed and therefore this configuration provides a good thermal insulation from the surrounding corona. But the question of the magnetohydrostatic equilibrium is still open. The current in the Kuperus-Raadu prominences flows in the direction opposite to that of the Kippenhahn-Schlüter type. It should be mentioned that in both configurations a field component along the prominence axis can be added without changing the models, and the observations indicate that this component is usually very large. In these cases Figs. 1a and 1b represent the projection into the plane of an actually 3-dimensional field. In the following we shall address the problem of prominence support in the Kuperus-Raadu model. Recently Malherbe and Priest (1983) and also Pneuman (1983) have dealt with this question. Malherbe and Priest take the prominence as an infinitely thin current sheet and use the theory of complex functions to describe the field. They present several possible configurations of the Kuperus-Raadu type. One class of their solutions has the property that all field lines which penetrate the prominence extend to very large heights, which would imply that these prominences are suspended from above. This seems to be a very unlikely situation, moreover it is difficult to imagine how in such a case a large field component along the axis of the prominence could be produced. They also present a configuration which is similar to that of Fig. 1b. This configuration has an upward directed Lorentz force everywhere in the prominence except at the upper edge. At this edge they have a line current with an infinite current density, resulting in an infinite downward Lorentz force density. Therefore this configuration is locally not in equilibrium, but more seriously even globally it cannot support a prominence because it turns out that the integrated upward Lorentz force is less than the downward force of the line current. Pneuman in this investigation argues that configurations such as in Fig. 1b should arise in the solar corona due to field line reconnection but he does not give an explicit representation of such a magnetic field. For these reasons one must consider the question of prominence support in Kuperus-Raadu configurations as still being open.

In the remainder of our discussion we shall describe some problems arising from such field configurations with the help of simple current systems.

1) Equilibria with line currents

We shall represent the coronal field above a bipolar photospheric region by a (fictitious) line current with strength I_0 placed at $x = 0$, $z = -z_0$ (the photosphere is taken as $z = 0$), shown in Fig. 2a). The total prominence current may be I_1 . As a

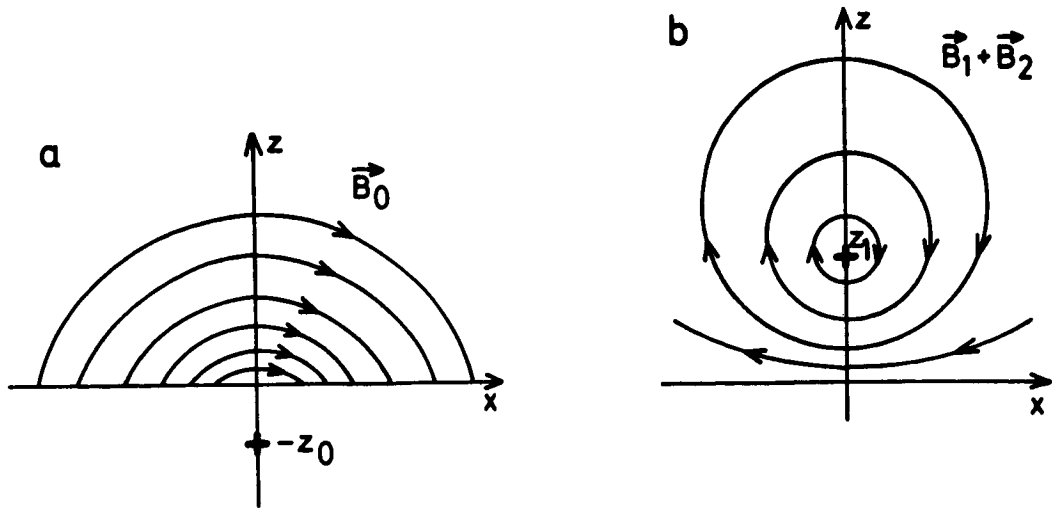


Fig. 2: a) Magnetic field of a bipolar region,
b) field which results from a line current at z_1 and its
subphotospheric mirror current.

first approximation we take it to be a line current at $x = 0$, $z = z_1$ and with the same direction as I_0 . Because of its high conductivity and rigidity, the photosphere acts as a mirror for I_1 (see also Kuperus and Raadu). Then the induced photospheric currents can be represented by a mirror current $-I_1$ at $z = -z_1$. The field resulting from I_1 and its mirror image ($\vec{B}_1 + \vec{B}_2$) is shown in Fig. 2b. The total force on the line current then is given by

$$F_z = I_1 \left(\frac{I_1}{2z_1} - \frac{I_0}{z_0 + z_1} \right). \quad (1)$$

The prominence can only be supported if $F_z > 0$ holds, which leads to

$$\frac{I_1}{2z_1} > \frac{I_0}{z_0 + z_1}. \quad (2)$$

Therefore prominence support is in principle possible, provided the current in the prominence is large enough. The essential question now is: what can produce this current I_1 ? If one piles heavy material onto the tops of the fieldlines of \vec{B}_0 then the field will sag down and lead to a configuration like in Fig. 1a with a current which is opposite to the required I_1 . One therefore has to conclude that only an upward stretching of \vec{B}_0 can give the right orientation of the current. This stretching could result from the action of the solar wind. If reconnection occurs in this situation one could obtain a field of the type shown in Fig. 3. It is characterized

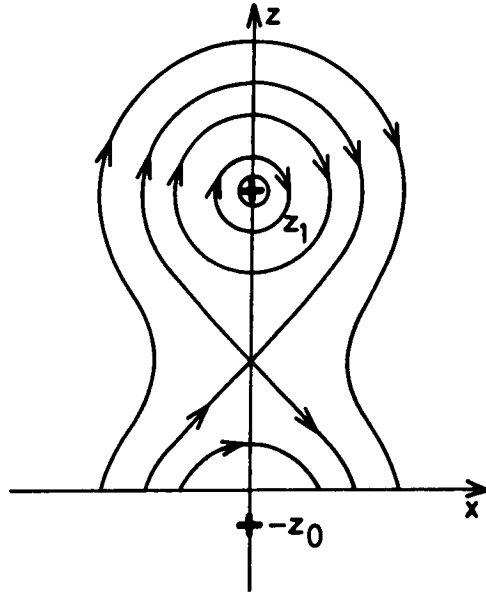


Fig. 3: Combined field of bipolar photospheric region and line current at z_1 .

by line currents I_0 at $-z_0$ and I_1 at z_1 . The effect of the solar wind on the current system now is reduced and the magnetic force is given by

$$F_z = - \frac{I_1 I_0}{z_1 + z_0} . \quad (3)$$

The line current I_1 will therefore move downward. A displacement of I_1 from z_1 to z_1^* can most easily be described by subtracting I_1 at z_1 and adding it at z_1^* to the initial field configuration. Because of the properties of the photosphere one also has to include the mirror currents I_1 at $-z_1$ and $-I_1$ at $-z_1^*$. From this then one obtains the field at the new position of the current z_1^* ,

$$B_x = \frac{I_0}{z_1^* + z_0} + \frac{I_1}{z_1^* + z_1} - \frac{I_1}{2z_1^*} . \quad (4)$$

With $F_z = -I_1 B_x$ the condition for support then gives

$$\frac{I_0}{z_1^* + z_0} + \frac{I_1}{z_1^* + z_1} < \frac{I_1}{2z_1^*} . \quad (5)$$

The photospheric magnetic fields are much larger than prominence fields, therefore in general I_0 will also be much larger than I_1 . But for simplicity let us take $I_1/I_0 = 1/2$. Then one finds that for $z_1/z_0 = 2$, $z_1^*/z_0 < 0.24$ must hold and for $z_1/z_0 = 8$, $z_1^*/z_0 < 0.30$ in order to give support. Now the length z_0 is typically the scale of the bipolar magnetic region in the photosphere, and if we take it to be ≈ 30000 km then $z_1^* < 10000$ km is required. This represents an extremely low

height for quiescent prominences. Therefore we conclude from our model calculations that in principle a global support in the Kuperus-Raadu case is possible, but we face the difficulty that these prominences would have to lie much lower than is usually observed.

2) Current sheets

The representation of prominences by line currents can only give answers to the questions of the global behaviour of the system, but not about the structure of the prominence itself. Quiescent prominences are known to be thin structures which have large vertical extent. Therefore they are more appropriately described by a current sheet. We shall use a very simple sheet extending from $z = a$ to $z = b$ having the following current distribution

$$j = I_1 \frac{6}{(b-a)^3} \frac{(b-z)(z-a)}{1} \quad (6)$$

which is normalized to

$$\int_a^b j(z') dz' = I_1 .$$

From

$$B_x = - \int_a^b \frac{j(z')(z'-z)}{x^2 + (z'-z)^2} dz' \quad (7)$$

one then obtains for $x \rightarrow 0$

$$B_x(0, z) = -I_1 \frac{6}{b-a} \left(\frac{1}{2} \frac{b+a-2z}{b-a} + \frac{(b-z)(z-a)}{(b-a)^2} \ln \left| \frac{b-z}{a-z} \right| \right) . \quad (8)$$

The function in brackets has a maximum of about 0.6 and therefore

$$B_x(0, z) \Big|_{\max} \approx 3.6 \frac{I_1}{b-a} \quad (9)$$

holds.

One also finds that far away from the sheet (i.e. for $r > b-a$) the field is very close to that of a line current I_1 . Therefore we shall describe the mirror effects of the photosphere by line currents of strength $\pm I_1$, at $z = -z_1$ and $z = -z_1^*$ respectively (as in the previous section). The resulting field then is again the one given by Eqn. (4), but in addition we now have to include the field of the current sheet. With $a = z_1^* - \delta$ and $b = z_1^* + \delta$ we then obtain

$$B_x(\text{sheet}) = -\frac{3I_1}{\delta} \left(\frac{z_1^*-z}{\delta} + \frac{z_1^*+\delta-z}{2\delta} - \frac{z-z_1^*+\delta}{2\delta} \ln \left| \frac{z_1^*+\delta-z}{z_1^*-\delta-z} \right| \right). \quad (10)$$

In this case now the condition for prominence support is that $F_z > 0$ must hold everywhere in the sheet (i.e. for $z_1^*-\delta < z < z_1^*+\delta$) and since the maximum B_x produced by the sheet is $1.8 I_1/\delta$ one has the new condition

$$-\frac{I_0}{z_1^* + z_0} - \frac{I_1}{z_1^* + z_1} + \frac{I_1}{2z_1^*} > \frac{1.8 I_1}{\delta}. \quad (11)$$

From this then one easily obtains a simple necessary (but not sufficient) condition

$$\frac{I_1}{2z_1^*} > \frac{1.8 I_1}{\delta}. \quad (12)$$

Since the whole prominence must lie above the surface of the sun $z_1^* > \delta$ must hold, and therefore condition (12) cannot be fulfilled. Physically this means that in such a configuration the upper parts of the prominence will be pulled downward and the prominence cannot exist as an extended vertical sheet. At present there seem to exist no models which can explain extended quiescent prominences in the framework of the Kuperus-Raadu configuration.

Acknowledgements

Part of this work was done during a visit to the University of St. Andrews, Scotland, and I would like to thank Dr. Eric Priest for his hospitality. I also thank Dr. J.L. Leroy for communicating his observational results prior to publication.

References

- Athay, R.G., Querfeld, C.W., Smartt, R.N., Landi degl'Innocenti, E., Bommier, V.: 1983, *Solar Phys.* 89, 3
 Kippenhahn, R., Schlüter, A.: 1957, *Z. Astrophys.* 43, 36
 Kuperus, M., Raadu, M.A.: 1974, *Astron. Astrophys.* 31, 189
 Leroy, J.L.: 1984, *privat communication*
 Leroy, J.L., Bommier, V., Sahal-Brechot, S.: 1983, *Solar Phys.* 83, 135
 Malherbe, J.M., Priest, E.R.: 1983, *Astron. Astrophys.* 123, 80
 Sahal-Brechot, S., Bommier, V., Leroy, J.L.: 1977, *Astron. Astrophys.* 59, 223

SECTION 2

**TECHNIQUES FOR MEASUREMENTS OF
VECTOR FIELDS**

TRENDS IN MEASUREMENT OF SOLAR VECTOR MAGNETIC FIELDS USING THE ZEEMAN EFFECT

J. HARVEY

National Solar Observatory, National Optical Astronomy Observatories,
Tucson, AZ 85726, U.S.A.*

Abstract. The goal of magnetic field observers is to measure $\mathbf{B}(r, \theta, \phi, t)$ over as large a range of the variables and with as much resolution as possible. The Zeeman effect provides one way of making such measurements and is the basis for most of our knowledge about the solar magnetic field. Observation of the Zeeman effect requires high precision spectropolarimetry to define the Stokes vector $\{I, Q, U, V\}$ across one or more spatial, spectral and temporal elements. During the past 70 years there has been much progress in methods for observing the Stokes vector. This article is a discussion of recent trends in spectropolarimetry as applied to the problem of Zeeman effect measurement. Difficulties of interpreting such measurements in terms of magnetic fields are not discussed.

The major trend is using detector arrays to improve observing efficiency. This has required new polarization modulation schemes that match the time required to read detector arrays. Another significant trend is away from the use of grating spectrographs as spectral isolation devices to narrow-band filters, on the one hand, to improve angular and temporal coverage, and to Fourier transform spectrometers, on the other hand, to improve spectral coverage and precision. Telescopes have come under scrutiny with the result that low-polarization designs and improved methods for compensating instrumental polarization have been developed. A requirement for high angular resolution suggests using adaptive optical devices to subdue the effects of bad seeing. The ultimate strategy to beat the seeing is to loft the telescope above the atmosphere such as is planned with a 30-cm telescope in 1985 and a 125-cm telescope in 1990.

1. Introduction

The goal of solar magnetic field observers is easily stated: Measure $\mathbf{B}(r, \theta, \phi, t)$ where \mathbf{B} is the magnetic field vector, r, θ, ϕ are coordinates in a spherical system and t is time. Achieving this goal is not easy. *In situ* measurements are restricted to large values of r , a small range of θ and very sparse sampling in coordinate space. Remote measurements are restricted by the requirement to observe more or less indirect effects of \mathbf{B} on electromagnetic radiation and by various limits in the resolution and span in coordinate space. Electromagnetic radiation from an object is affected in many ways by a magnetic field. Therein lies both opportunity and difficulty. A particular effect may give us the opportunity to remotely measure a magnetic field but other effects, operating at the same time, often corrupt the clean result one seeks. One of the cleanest effects of a magnetic field on electromagnetic radiation is the Zeeman effect in the formation of spectrum lines. This brief report concerns recent trends and developments in the use of the Zeeman effect to obtain measurements of \mathbf{B} . The emphasis is on instrumental

*Operated by the Association of Universities for Research in Astronomy, Inc., under contract with the National Science Foundation.

and observational matters since interpretational problems are dealt with elsewhere in these proceedings [see the reviews by Stenflo (1978) and Rees (1982)].

Neglecting problems of interpretation, the measurement of \mathbf{B} using the Zeeman effect reduces to observing the Stokes polarization vector $\{I, Q, U, V\}$ as a function of wavelength (λ) across one or more suitable spectrum lines. This is an old problem in solar physics which dates back some 70 years. Progress has been episodic, following major instrumental advances every solar cycle or so. In spite of this long heritage, an ideal method of observing \mathbf{B} by using the Zeeman effect still eludes us. It is one of the most difficult instrumental challenges in astrophysics.

An excellent review of the techniques of optical solar polarimetry in use as of 1980 was prepared by Baur (1980). I do not intend to repeat his work but only to discuss some recent developments and trends for the near future.

2. Requirements

2.1. ANGULAR RESOLUTION AND SPAN

Since the solar magnetic field is highly intermittent, very high angular resolution is necessary to avoid an unrecoverable loss of information caused by angular integration. A fundamental observational limit is set by the photon mean free path in the photosphere of about 100 km. Though it is likely that smaller-scale structures exist in the magnetic field, they will appear as if in a fog with a characteristic scale of 100 km or so. Thus, a goal for angular resolution is to be able to measure low-contrast features at a scale of about 0.1 arc sec. Many of the targets of \mathbf{B} observations involve active regions with a typical scale of 3 arc min which thus becomes a reasonable goal for the angular span of \mathbf{B} observations.

2.2. SPECTRAL RESOLUTION AND SPAN

One criterion might be that spectral resolution should be sufficient to allow one to unambiguously fit a model line profile to observations. Another criterion might be that resolution should be adequate to faithfully represent a spectral line profile. This second criterion is more conservative but more stringent than the former. Where to set a requirement depends on one's faith in modeling and restoration procedures. My own approach is to reduce reliance on modeling as much as possible; hence, I favor the high resolution criterion. Experience with the Fourier transform spectrometer at the McMath telescope, which has superb spectral resolution, indicates that the spectral modulation transfer function should be nearly unity up to 60 cycles per Å at 6000Å. This may seem excessive, but recent observations by Stenflo *et al.* (1984) and Stenflo and Harvey (1984) reveal significant errors if resolution of this order is not used.

The spectral span requirement follows from a need to observe several different heights in the solar atmosphere to define the pattern of \mathbf{B} as a function of r . The deepest spectral lines are formed near the 1.6μm wavelength of the minimum of continuous spectral opacity. The highest lines are formed in the transition region and corona and are accessible only in the ultraviolet. Thus a large span of wavelength is required to cover the greatest possible height range. A goal might be to try to span 6 decades of density by using 6 different spectral lines. (See also Lozitskii and Dolgoplov 1983).

2.3. TEMPORAL RESOLUTION

Large Doppler motions of the photosphere result from the acoustic oscillation modes trapped underneath. These motions produce a blurring of spectral line profiles if the observational time resolution is not adequate to resolve the motion. Fortunately, these motions are very nearly in phase and radial at

heights between the low photosphere and the low chromosphere, so observations at different heights are disturbed in similar ways. Thus we are faced with essentially a fluctuating wavelength zero point. This is only a problem if the time resolution for a Zeeman measurement is inadequate to avoid spectral blurring of the line profile (especially if such blurring is different for the different elements of the Stokes vector). Potentially more serious is the possibility of distorting effects of high frequency, propagating waves on Zeeman profile measurements. In the chromosphere, these waves have significant amplitudes at periods as short as 30 seconds which indicates a requirement for a Zeeman measurement to be completed in a matter of a few seconds in order to freeze the distortion of the spectral line profile. Another criterion is the sound travel time across an angular resolution element. Both of these criteria indicate that sampling should be completed in ~ 10 sec. In order to resolve time-varying activity in an active region, the entire region should be observed within the sampling period. The length of a sequence of Zeeman measurements should be great enough to allow the Zeeman and Doppler effects on line profiles to be separated.

2.4. SENSITIVITY

The amount and state of polarization produced by the Zeeman effect across a spectrum line depends on \mathbf{B} . A reasonable goal might be to define \mathbf{B} with an amplitude error of ± 10 G and a direction error of ± 0.1 radian. However, an important research topic is the nature of the electric current, \mathbf{J} , in the solar atmosphere and since $\mathbf{J} = \nabla \times \mathbf{B}$ the observational accuracy required of \mathbf{B} may well be more stringent to obtain a useful value of \mathbf{J} . Modeling of the Zeeman effect in a weak field shows that

$$\begin{aligned} Q(\lambda) &= C_Q B^2 \sin^2 \theta \cos 2\phi \frac{\partial^2 I(\lambda)}{\partial \lambda^2} \\ U(\lambda) &= C_U B^2 \sin^2 \theta \sin 2\phi \frac{\partial^2 I(\lambda)}{\partial \lambda^2} \\ V(\lambda) &= C_V B \cos \theta \frac{\partial I(\lambda)}{\partial \lambda} \end{aligned}$$

where θ and ϕ are polar coordinates in a system whose origin is at the observed source and whose pole points to the observer, and the C 's are constants. For a typical atomic spectrum line in the visible, the peak amplitude of Q , U , and V can be computed, and assumed errors of measurement converted to errors in B , θ and ϕ . When this is done, one finds that the polarimetric precision required to achieve the 10 G and 0.1 radian levels is of the order of 10^{-4} I or less. This small level is very stringent and indicates why vector field measurement is so difficult.

2.5. SUMMARY AND IMPLICATIONS

The requirements discussed above are summarized in Table 1.

Taken together, the requirements in Table 1 imply a need for a large telescope having an aperture in excess of 1 m just to obtain the required angular resolution. A large aperture is also required to collect enough light. For example, to collect $\sim 10^8$ photons from $(0.05 \text{ arc sec})^2$ and $20 \text{ m}\text{\AA}$ in 10 seconds implies an aperture of at least 1.4 m at 5000\AA . At other wavelengths, the required aperture is even larger. Another severe implication is a stupendous data rate of the order of 1 Gbyte per second.

Needless to say, no existing or planned instrument meets the needs outlined here. Various practical compromises have been made in the effort to achieve a useful vector magnetograph. The major compromises for ground based instruments have been to restrict the angular resolution to a value similar to seeing limitations and to observe a single spectral line.

Table 1. Requirements for observation of $I(x,y,\lambda,t)$.

Parameter	Element	Rationale	Number	Rationale	Implication
Angular (x,y)	0.05 arc sec	scale height in photosphere	4000x3000	cover an active region	>1m aperture telescope
Spectral (λ)	0.02Å @ 6000Å	resolve spectral structure	20 per line × 6 lines	cover a line cover a height range	large wavelength range spectrometer
Temporal (t)	10 sec	sound travel time across Δx or Δy	1	resolve oscillations and activity of active region	high data rate
Sensitivity for I , Q/I , U/I , V/I	$\sim 10^{-4}$	define B adequately	$\sim 10^4$ or 2 bytes	adequate dynamic range	very high data rate

3. Recent progress toward Stokes polarimeter facilities

3.1. IMPROVEMENTS

Several older facilities have been upgraded for polarimetry recently. The Stokes polarimeter at the Mees observatory now incorporates the "Stokes II" detector system formerly operated by the High Altitude Observatory at Sacramento Peak (Baur *et al.* 1981; McCabe 1983; Mickey, these proceedings). One of the early results is reported to be the first full disk raster of the sun in all 4 Stokes parameters across the 6303 Å line. The resolution was 30", and 90 minutes were required to complete the observation.

A new CCD camera and data processing system have been deployed at the Marshall Space Flight Center vector magnetograph (West *et al.* 1983; Hagyard *et al.* 1984; West, these proceedings).

The Arosa station of the ETH Institut für Astronomie has been upgraded for polarimetry (Stenflo 1983).

The photoelectric Stokes polarimeter at Okayama is now in frequent operation with careful compensation for instrumental polarization (Makita *et al.* 1982; Sakurai and Makita 1984).

At Kitt Peak, the Stokes polarimeter (Harvey *et al.* 1980), operated with the 1-m Fourier transform spectrometer (Brault 1978), is complete except for the addition of a means to measure both linear polarization parameters simultaneously. An achromatic telescope polarization compensator is now available.

3.2. FACILITIES AND IMPROVEMENTS IN PREPARATION

The most exciting near-term development is the upcoming flight of Spacelab 2 scheduled for April 1985. If all works well, the Solar Optical Universal Polarimeter will provide 1/3" Stokes images of the sun across a number of spectrum lines over a period of several days. Additional flights in 1986 and 1987

may take place.

A major instrument (THEMIS) is under development by French astronomers to be located in the excellent observing conditions found on the Canary Islands (Rayrole, 1982; Rayrole and Ribes, 1984). This 90-cm diameter telescope is intended to provide excellent quality images free of polarization effects to a spectrograph designed to record all the Stokes parameters across a line profile.

A video vector magnetograph is under development in China for the Purple Mountain Observatory (Ai and Hu 1981).

Although vacuum telescopes have been a popular way to reduce seeing degradation, the windows are subjected to great stress by the pressure of the atmosphere. This can produce nonuniform birefringence in the windows which greatly complicates Stokes polarimetry. One proposed solution is to make the windows thinner and replace the vacuum with helium at or near the atmospheric pressure. A plan to do this to the Tower telescope at Sacramento Peak has been proposed by R. Dunn. A Stokes polarimeter would then be added at the focal plane.

A video processor and polarization analysis system is planned for the San Fernando Observatory (Richter and Zeldin, these proceedings).

3.3. LONG RANGE PLANS

Plans are well defined for constructing a Solar Optical Telescope to operate in space starting in 1990. The project has not progressed to the detailed design stage yet, but one hopes that this will occur in 1985. Current plans provide for Stokes polarimetry across a number of spectrum lines with an angular resolution of about 0.1 arc sec and adequate spectral and temporal resolution to attack many difficult problems.

The Joint Organization for Solar Observations (JOSO) hopes to build a Large European Solar Telescope. Published plans (Wyller 1983) indicate an aperture of $\sim 2.5\text{m}$ and a design with low intrinsic polarization that would permit accurate Stokes polarimetry.

A consortium of the University of Hawaii, Sydney University, the High Altitude Observatory and the National Solar Observatory have discussed building a new Stokes polarimeter and telescope. At present, plans are in a preliminary state.

There has been discussion of a large solar telescope (2.1m aperture) to be built in the Soviet Union (Grigoryev and Karpinsky, 1981; Karpinsky 1982). This instrument would include a Stokes polarimeter.

4. Recent progress regarding elements of vector magnetographs

4.1. SYSTEM DESIGN

Cox (1983) has applied the concept of information capacity to assess the relative performance of different polarimeter designs. The information capacity per measurement is $R(\psi) = 1/2 \log_{10}[(\psi^2 + \sigma_\psi^2)/\sigma_\psi^2]$ where ψ is one of the components of a Stokes vector and σ_ψ is its standard deviation. Using this criterion, Cox concludes that a photoelastic modulator is superior to a rotating analyzer for simultaneous measurement of Q , U and V , provided that a dual beam system of the sort used by Eckstrom *et al.* (1982) is employed.

A difficult problem is to acquire polarization, spectral and spatial information, simultaneously (or at least rapidly) at a large number of points. Elements of this old problem have been recently discussed by Atherton (1983) and Girard (1984). No practical solution that avoids scanning in some domain

seems to have been discovered.

4.2. EXOTIC ASPECTS OF THE ZEEMAN EFFECT

A large number of atomic spectrum lines have recently been observed (Chang and Noyes 1983) at wavelengths around $13\mu\text{m}$. Because the Zeeman effect varies as λ^2 and line width varies only as λ , these spectrum lines are far more sensitive than visible lines to magnetic fields (Brault and Noyes 1983). Their potential for diagnostic purposes is offset by several problems: They are formed fairly high in the atmosphere. Diffraction at $13\mu\text{m}$ imposes a requirement for a gargantuan telescope aperture to attain 0.1 arc sec angular resolution. Detector technology is not very advanced for work at $13\mu\text{m}$. Telescope components (unless cryogenically cooled) glow at $13\mu\text{m}$ so that unusual care is required to attain good signal to noise ratios. For these and other reasons, it does not seem likely that Stokes polarimetry will be practised routinely in the near future at $13\mu\text{m}$.

A few atomic spectral lines have been identified recently (Stenflo *et al.* 1984) that exhibit negative Zeeman splitting. Stokes polarimetry with a pair of negative and positive Zeeman split lines may offer some advantages for assessing instrumental effects or magnetooptic effects in the solar atmosphere.

The Zeeman effect in the spectra of several diatomic molecules is interesting. In addition to cases of positive and negative Zeeman splitting in the same band (e.g. the OH 2-0 band illustrated in Figure 1), strong intensity asymmetries of the Stokes parameters exist (Schadee 1978) across many molecular spectrum lines. This suggests the possibility of an independent check of results obtained by classical analyses of atomic Zeeman splitting.

Asymmetries of the Stokes profiles of atomic lines are also observed (Stenflo *et al.* 1984; Stenflo and Harvey 1984) but in this case, the cause seems to be due to height gradients of both **B** and mass motions within the magnetic field structure. Such asymmetries result in net polarizations from magnetic features in broad wavelength observations (Kemp and Henson 1983; Henson and Kemp 1984) which prompted Stenflo (1984) to suggest exploiting this to measure the line-of-sight component of a magnetic field. If we combine this idea with the broadband transverse field magnetograph (Leroy 1962), we obtain the concept of a broadband vector magnetograph. The probable difficulties of calibration of such an instrument make it possibly of limited interest.

4.3. ANGULAR RESOLUTION

Plans for large telescopes have been mentioned above. Short of going to space, achieving high angular resolution is a matter of careful telescope design and construction and finding a site with excellent seeing. The free atmosphere at good sites will permit occasional glimpses of sub-arc-second resolution, but these occasions cannot be provided on demand. In principle, it is possible to ameliorate the effects of bad seeing by sensing and correcting the wavefront distortions produced by the atmosphere. Several active optical systems intended to offset bad seeing have been proposed and built (Hardy 1982; Smithson *et al.* 1983). To date none has succeeded in producing a near-diffraction-limited image of the sun. It is reasonable to expect eventual success since there is no known insurmountable technical barrier. However, a serious limitation will be the very small field (a few arc seconds) that can be restored. As a more modest goal, image stabilizing systems can double the resolution of solar images and should, in my opinion, be a part of every solar telescope.

4.4. TELESCOPE POLARIZATION

Telescopes alter the state of polarization of light. This can be conveniently described by a 4x4 Mueller matrix **M** that operates on the input Stokes vector: $I_{\text{out}} = \mathbf{M}I_{\text{in}}$. Ideally, **M** is a constant scalar

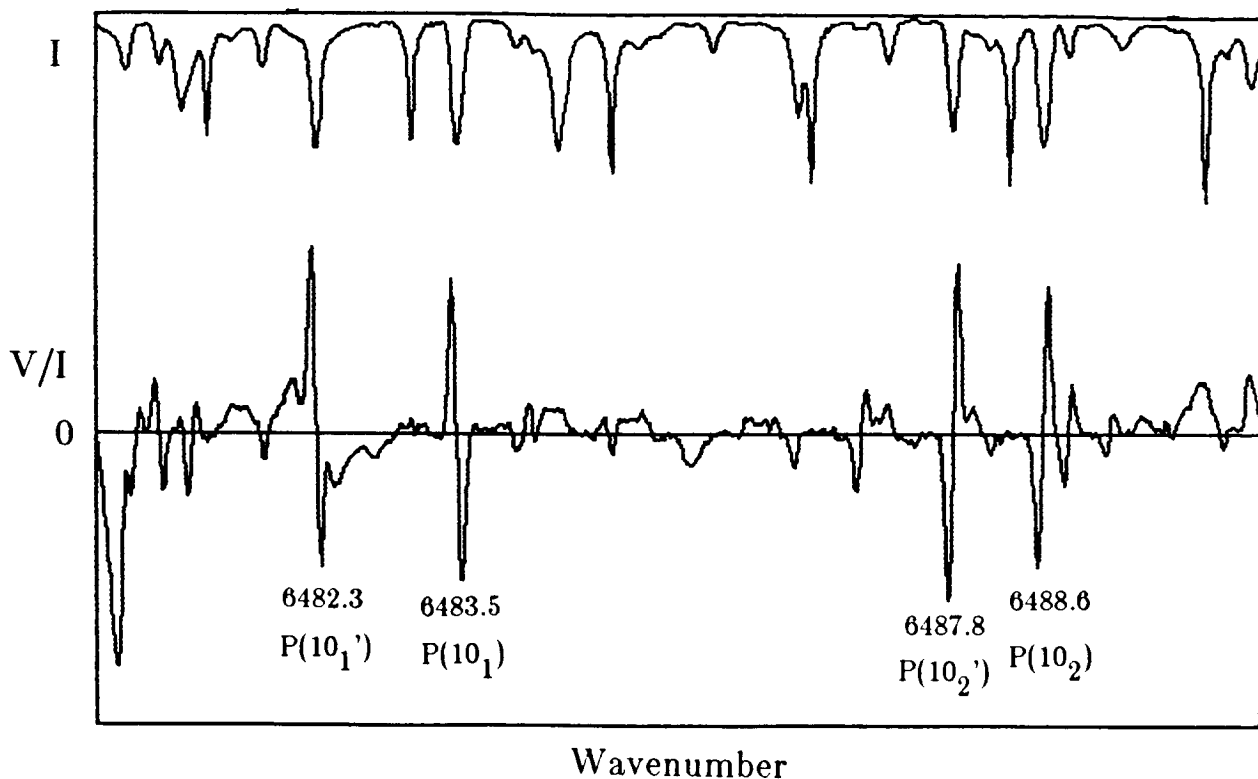


Fig. 1. A portion of the infrared spectrum of a sunspot umbra covering 6480 to 6490 cm^{-1} observed with the FTS polarimeter at Kitt Peak. The upper trace is the intensity in unpolarized light. The lower trace is the degree of circular polarization on an expanded ordinate scale. Notice that the left-hand pair of OH 2-0 lines have opposite sense of Zeeman splitting to that of the right-hand pair.

matrix with unit diagonal elements and zero off-diagonal elements so that $I_{\text{out}} = I_{\text{in}}$. In practice, \mathbf{M} is not so nice. There is a trend toward using symmetric optical systems such as Cassegrain and Gregorian systems to minimize polarization effects. Unfortunately, many telescopes now used for polarimetry involve oblique, polarizing reflections. For example, an \mathbf{M} for the McMath telescope for $\lambda = 5250\text{\AA}$, a declination of 20° and an hour angle of 45° is

$$\mathbf{M} = \begin{bmatrix} .775 & .035 & .008 & .003 \\ .035 & .771 & .028 & -.069 \\ .009 & .001 & .715 & .297 \\ .003 & .075 & -.295 & .712 \end{bmatrix}$$

This implies (first column) a serious conversion of unpolarized input light into a polarized output and conversion of one form of polarization into another (columns 2-4). The problem has recently been discussed by Bachmann *et al.* 1975; Nikonov and Nikonova 1982; Makita *et al.* 1982; and Nariai, 1982. A frequent approach is to provide an optical compensation consisting of tipped glass plates and fractional waveplates.

In the case of the McMath Fourier transform spectrometer polarimeter, a need for broadband compensation prevented the use of tipped glass and transmissive waveplates. (Previous experience with these was also unsatisfactory because of variable heat-induced birefringence in the solar beam). The old idea of compensating the polarizing effects of oblique reflections by additional oblique reflections

(e.g. Smith 1956) was used. The constantly changing geometry of the telescope led to a rather complicated mechanical mounting which requires adjustment only for different declinations (Figure 2). Six reflections were required which reduces light flux substantially. By far the better solution is to build a low-polarization telescope at the start.

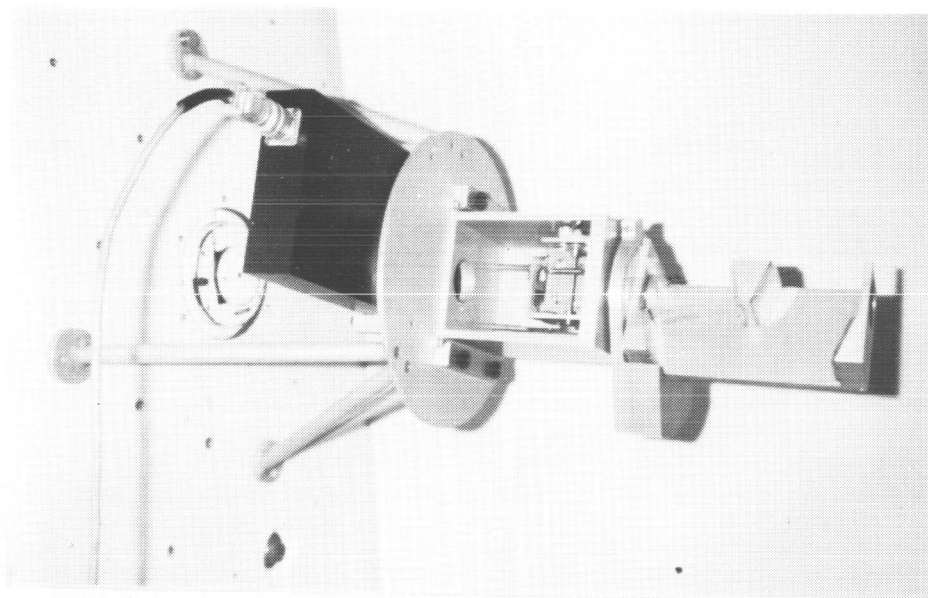


Fig. 2. Front end of the Kitt Peak FTS polarimeter, sunlight proceeds from right to left. The first unit consists of four mirrors mounted in such a way as to produce an on-axis beam that is compensated for the polarization produced by the final mirror in the telescope. This unit is mounted on a bearing and weighted to maintain a fixed position as the rest of the system rotates. The next unit contains two mirrors to compensate for the main heliostat mirror of the telescope. An adjustment for changing declination of the sun is provided. The final unit (black box) contains a photoelastic modulator and Glan polarizer.

With the advent of ultrahigh reflectivity mirror coatings, one might think that instrumental polarization is an obsolete problem. Unfortunately, many such coatings exhibit strong retardation between s and p wave reflections (except for limited wavelength and inclination ranges), and the need for compensation is still with us.

4.5. POLARIZATION MODULATORS

Dozens of methods for measuring the state of polarization have been proposed. Beam division is one popular approach. Here a beam splitter of some sort produces two or more images that contain various combinations of the Stokes parameters. Azzam (1982) describes a recent 4-beam polarimeter. A major problem with all such techniques is sensitivity to drifts in the detectors and the beam splitter itself which limit precision to not much better than 1%. A three-beam polarimeter in use at the National Solar Observatory and a result are shown in Figures 3 and 4.

Sensitive polarimetry requires modulation of the beam in time to reduce errors due to instrumental drifts. A common approach is to use one or more electrically-variable retardation devices. Pockels cells are frequently used but suffer from a narrow field of view that restricts the f/number and requires

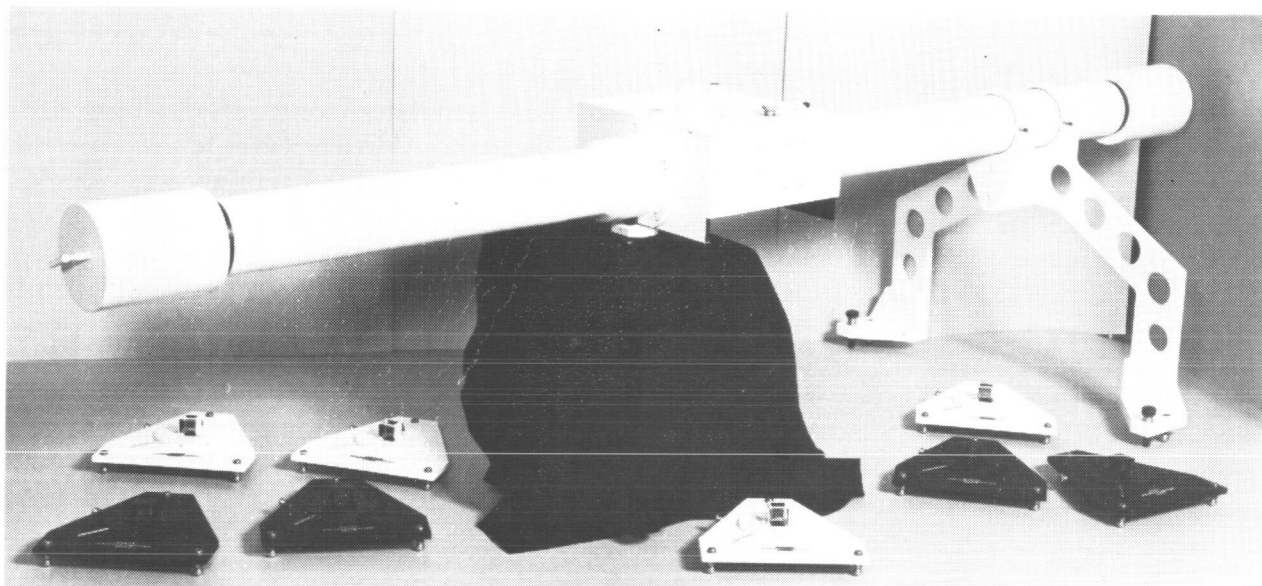


Fig. 3. A multiple beam polarimeter provided for the spectrograph at Kitt Peak. The input solar image is divided by the large structure into three beams which undergo 0 , $1/4\lambda$ and $1/2\lambda$ retardations. These beams proceed to one of several linearly polarizing beam splitters to produce six beams simultaneously on the entrance slit of the spectrograph.

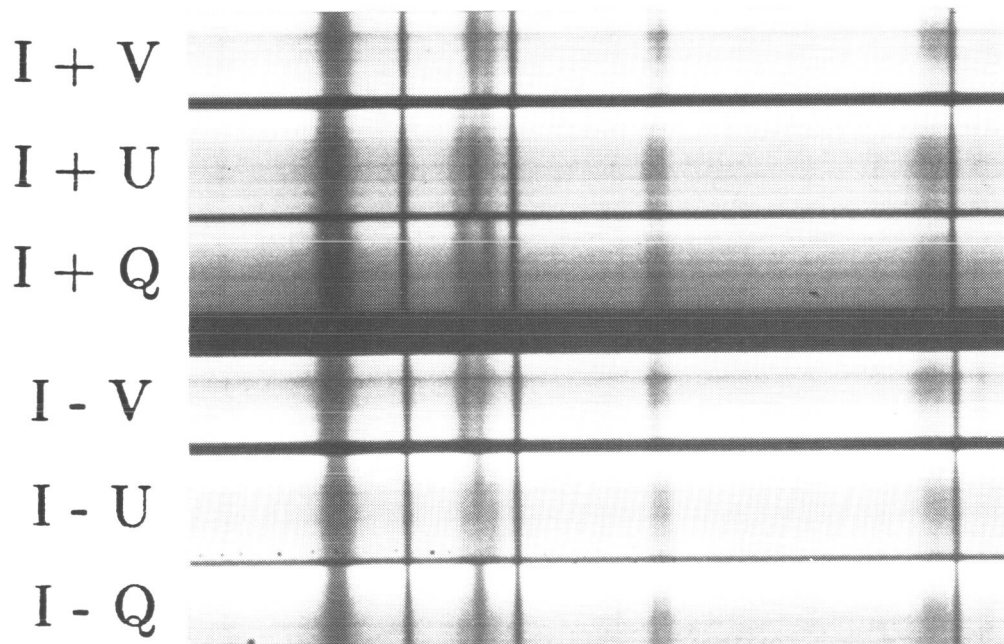


Fig. 4. Spectra of the umbra of a small sunspot produced by the polarimeter shown in Figure 3. The spectral region shown is near 6303\AA .

careful alignment to avoid introducing systematic measurement errors (Ai and Hu, 1979; Grigoryev and Ilgamov, 1983). The field can be widened with compensating plates of MgF (West 1978) or by using two Pockels cells separated by a $\lambda/4$ retarder (Ai and Hu, 1981). Additional problems include temperature sensitivity and electrode degradation.

These difficulties made the introduction of photoelastic modulators welcome indeed (see review by Kemp 1981). The basic idea is to introduce time-varying strain birefringence in an otherwise non-birefringent medium such as fused quartz. These devices have produced the highest precision solar polarimetry (Kemp, 1982). One problem is the high modulation frequency of these devices (10's of kHz) which makes their use with area detectors difficult. One solution is to use the beat between the frequencies of two nearly identical photoelectric modulators to produce a low-frequency modulation (Kemp *et al.* 1977, Wolstencroft 1978, McLean 1980, Stenflo 1984). As pointed out by Cox (1983), this technique suffers from the light loss at 4 surfaces and is a rather low-efficiency modulation. Cox mentions the possibility of antireflection coatings to reduce reflection losses, but he also notes that the process of producing such coatings often induces stress (and thus birefringence) within the material. A better solution may be to stress the material directly at the desired modulation frequency. Such a modulator was developed by Angel (Stockman, 1982).

Other modulators have been discussed. The Kerr effect in nitrobenzene is used at Kitt Peak to modulate light for a magnetograph (Livingston *et al.* 1976). We find that it is necessary to replace the liquid after about 2 years of service. Otherwise, the device has been satisfactory. The analogue of the Kerr effect in a transparent ceramic material (PLZT) can be used as a polarization modulator. Experience at Kitt Peak has not been encouraging owing to scattered light in the material, a high temperature coefficient and hysteresis.

Nematic liquid crystal films offer promise as low-frequency polarization modulators (Kahn 1972; Armitage 1980; Jacobs 1981; Spillman and Soref 1982). These devices are under investigation by Rust (these proceedings) for use in solar polarimetry. Possible difficulties that one might expect in simple devices include a narrow field angle, slow response time, hysteresis and high temperature sensitivity.

4.6. SPECTROMETERS

Traditionally, grating spectrographs are used for observations of the Zeeman effect. Severe limitations of spectral and angular field coverage have stimulated the use of other spectrometers. In particular, birefringent filters appear to be the favored choice for new polarimeters. Improvements by the Lockheed group (Title and Torgerson 1983) have made birefringent filters quite acceptable for Stokes polarimetry. One worries a bit about the accuracy with which line profiles can be reconstructed from filtergrams, but adequate techniques appear to be available (*e.g.* Caccin *et al.* 1983).

The Fourier transform spectrometer is well established as a means to obtain excellent line profile measurements. Polarimetry with an FTS is also available (Harvey *et al.* 1980; Lipp and Nafie 1984). The major disadvantages of FTS observing are the sacrifice of time resolution and angular coverage to obtain excellent spectral coverage and resolution. These limitations could be overcome in principle by the use of aliased undersampling and array detectors, although severe technical challenges remain.

In consideration of weight and volume limitations for space experiments, Fabry-Perot (Gillespie and Title 1981; Rust, these proceedings) and Michelson interferometers (Stenflo 1984) have been considered as spectrometers. So far only the longitudinal Zeeman effect is intended to be studied.

In the future, further development of acousto-optic-tunable filters (*e.g.* Yano and Watanabe 1976) and electro-optic-tunable filters (Pinnow *et al.* 1979) may provide enough spectral resolution to be useful for Stokes polarimetry.

4.7. DETECTORS

Silicon detector arrays (CCD, etc.) are preferred for new polarimeters (e.g. McLean 1980; Tyson and Lee 1981; Stockman 1982; West 1983; Hagyard *et al.* 1984). Arrays as large as 1024x1024 have been produced. The devices enjoy high quantum efficiency and low noise but some of them do suffer from a few problems. Perhaps the major one is interference fringing in monochromatic light. The thin silicon wafer acts like a bad Fabry-Perot etalon. Precision photometry thus requires great care in calibration and good instrumental stability. Another problem is that the time required to read all the pixels is sufficiently long that seeing effects may be a major source of noise.

5. Conclusion

While at present we lack the observational abilities to fully exploit the Zeeman effect as a way of determining the vector magnetic field, recent improvements and future prospects indicate that substantial improvements in our understanding of the solar magnetic field will soon be realized.

References

- Ai, G.-X. and Hu, Y.-F. 1979, *Pub. Beijing Astron. Obs.*, No. 3, 63.
Ai, G.-X. and Hu, Y.-F. 1981, *Acta Astrophys. Sinica* **1**, 284.
Armitage, D. 1980, *Appl. Opt.* **19**, 2235.
Atherton, P. D. 1983, *Proc. S.P.I.E.* **445**, 535.
Azzam, R. M. A. 1982, *Opt. Acta* **29**, 685.
Bachmann, G., Jäger, F. W., Künzel, H., Pflug, K. and Staude, J. 1975, *H. Hertz. Inst. Report*, No. 4.
Baur, T. G. 1981, *Opt. Eng.* **20**, 2.
Baur, T. G., Elmore, D. E., Lee, R. H., Querfeld, C. W. and Rogers, S. R. 1981, *Solar Phys.* **70**, 395.
Brault, J. W. 1978, *Osserv. Mem. Astrofis. Arcetri*, **108**, 33.
Brault, J. and Noyes, R. 1983 *Astrophys. J.* **269**, L61.
Caccin, B., Falciani, R., Roberti, G., Sambuco, A. M. and Smaldone, L. A. 1983 *Solar Phys.* **89**, 323.
Chang, E. S. and Noyes, R. W. 1983, *Astrophys. J.* **275**, L11.
Cox, L. J. 1983, *Mon. Not. Roy. Astron. Soc.* **203**, 571.
Ekstrom, P. A., Stokes, R. A. and Stokes, G. M. 1981, *Opt. Eng.* **20**, 19.
Gillespie, B. A. and Title, A. M. 1981, *Bul. Am. Astron. Soc.* **13**, 888.
Girard, A. 1984, *J. Optics* **15**, 7.
Grigoryev, V. M. and Ilgarnov, R. M. 1983, *Soln. Dann.*, No. 8, 69.
Grigoryev, V. M. and Karpinsky, V. N. 1981, *Izved. Geom. Aeron. Solar Phys.* **56**, 57.
Hagyard, M. J., Cumings, N. P. and West, E. A. 1984, *Proc. Kunming Workshop on Solar Phys.*, in press.
Hardy, J. H. 1982, *Proc. S.P.I.E.* **332**, 252.
Harvey, J., Brault, J., Stenflo, J. and Zwaan, C. 1980, *Bul. Am. Astron. Soc.* **12**, 476.
Henson, G. D. and Kemp, J. C. 1984, *Solar Phys.*, submitted.
Jacobs, S. D. 1981, *Proc. S.P.I.E.* **307**, 98.
Kahn, F. J. 1972, *Appl. Phys. Lett.* **20**, 199.
Karpinsky, V. N. 1982, *Izv. Pulkovo Obs.*, No. 200, 135.
Kemp, J. C. 1981, *Proc. S.P.I.E.* **307**, 83.
Kemp, J. C. 1982, *Smiths. Astrophys. Obs. Spec. Report 992* **1**, 191.
Kemp, J. C. and Henson, G. D. 1983, *Astrophys. J.* **266**, L69.
Kemp, J. C., Rieke, G. H., Lebofsky, M. J. and Coyne, G. V. 1977, *Astrophys. J.* **215**, L107.

- Leroy, J. 1962, *Ann. d'Astrophys.* **25**, 127.
- Lipp, E. D. and Nafie, L. A. 1984, *Appl. Spect.* **38**, 20.
- Livingston, W. C., Harvey, J., Slaughter, C. and Trumbo, D. 1976, *Appl. Opt.* **15**, 40.
- Lozitskii, V. G. and Dolgoplov, V. I. 1983, *Soln. Dann.*, No. 5, 71.
- Makita, M., Hamana, S., Kawakami, H. and Nishi, K. 1982, *Ann. Tokyo Astron. Obs.* **19**, 24.
- McCabe, M. 1983, *Bul. Am. Astron. Soc.* **15**, 202.
- McLean, I. S. 1980, *Ricerche Astron.* **10**, 71.
- Nariai, K. 1982, *Ann. Tokyo Astron. Obs.* **19**, 55.
- Nikonov, O. V. and Nikonova, E. S. 1982, *Izv. Pulkovo Obs.*, No. 199, 103.
- Pinnow, D. A., Abrams, R. L., Lotspeich, J. F., Henderson, D. M., Plant, T. K., Stephens, R. R. and Walker, C. M. 1979, *Appl. Phys. Lett.* **34**, 391.
- Rayrole, J. 1982, *JOSO Ann. Rep.*, 54.
- Rayrole, J. and Ribes, E. 1984, *Proc. Kunming Workshop on Solar Phys.*, in press.
- Rees, D. E. 1982, *Proc. Astr. Soc. Australia* **4**, 335.
- Sakurai, T. and Makita, M. 1984, *Vector Magnetograms of Solar Active Regions*, Tokyo Astron. Obs.
- Schadee, A. 1978, *J. Quant. Spect. Rad. Trans.* **19**, 517.
- Smith, E. v. P. 1956, *Astrophys. J.* **124**, 43.
- Smithson, R. C., Marshall, N. K., Pope, T. P., Sharbaugh, R. J. and Tarbell, T. D. 1983, *Bul. Am. Astron. Soc.* **15**, 717.
- Spillman, W. B., Jr. and Soref, R. A. 1982, *Appl. Opt.* **21**, 2696.
- Stenflo, J. O. 1978, *Rep. Prog. Phys.* **41**, 865.
- Stenflo, J. O. 1983, *Astron. Gessel.* **59**, 294.
- Stenflo, J. O. 1984, *Appl. Opt.*, submitted.
- Stenflo, J. O. and Harvey, J. W. 1984, *Solar Phys.*, submitted.
- Stenflo, J. O., Harvey, J. W., Brault, J. W. and Solanki, S. 1984, *Astron. Astrophys.* **131**, 333.
- Stockman, H. S. 1982, *Proc. S.P.I.E.* **331**, 76.
- Title, A. M. and Torgerson, D. D. 1983, *Bul. Am. Astron. Soc.* **15**, 709.
- Tyson, J. A. and Lee, R. W. 1981, *Proc. S.P.I.E.* **290**, 144.
- West, E. A. 1978, *Appl. Opt.* **17**, 3010.
- West, E. A. 1983, *Bul. Am. Astron. Soc.* **15**, 392.
- Wolstencroft, R. D. 1978, *Observatory* **98**, 99.
- Wyller, A. A. 1983, LEST Foundation Executive Summary.
- Yano, T. and Watanabe, A. 1976, *Appl. Opt.* **15**, 2250.

THE HANLE EFFECT APPLIED TO MAGNETIC FIELD MEASUREMENTS

J.L. Leroy, Pic du Midi Observatory

ABSTRACT

The Hanle effect is the modification by a local magnetic field of the polarization due to coherent scattering in spectral lines. One can understand it as resulting from the precession of a classical oscillator about the magnetic field direction. The sophisticated quantum-mechanical treatment, which is required to compute the polarization parameters of scattered light, has been successfully developed in the recent years.

The main features of the Hanle effect concerning magnetic field measurements are :

1 - A good sensitivity within the approximate range $0.1 B_{\text{typ}} - 10 B_{\text{typ}}$, where B_{typ} is the field strength yielding a Larmor period equal to the radiative lifetime. 2 - There is no Hanle effect for field vectors parallel to the exciting beam ; for other orientations of the magnetic field one can observe generally a decrease of the polarization degree and a rotation of the polarization direction. 3 - The Hanle effect refers essentially to the linear polarization in a spectral line : the Stokes parameter V is roughly one order of magnitude smaller than Q and U . 4 - Various points in the line profile are affected in the same way by the change of linear polarization so that polarization parameters can be measured on the integrated line profile.

Measuring the integrated values of Q and U in two spectral lines is enough to determine the vector field, except for the fundamental 180° ambiguity on the transverse field which cannot be avoided (at least for optically thin lines). The detection of small polarization amounts, down to the limit imposed by photon statistics, is a well known technique. On the other hand, extreme care is wanted to keep the instrumental linear polarization negligible. The comparison of results obtained with two different instruments has shown a reasonable agreement and, eventually, the field orientation has been determined with an accuracy better than 10° in solar prominences.

The diagnosis may be more complicated, yet very interesting, under several circumstances :

1 - For increasing densities collisional depolarization becomes important, which, in turn, yields a way for determining the density. 2 - The polarization parameters are deeply modified if the scattering takes place within an optically thick medium. 3 - Unresolved fine-structure fields have a well-defined influence, mainly reducing the rotation of the polarization direction. 4 - For "large fields" information on the field strength is lost but one can determine the direction of the projected field vector ; this is the situation met with the forbidden coronal lines.

THE HANLE EFFECT APPLIED TO MAGNETIC FIELD MEASUREMENTS

I INTRODUCTION

The Hanle effect is the modification by a local magnetic field of the polarization due to coherent scattering in spectral lines (Sahal-Bréchet, 1981 ; Stenflo, 1982).

The classical theory of the Hanle effect, which is sufficient for giving a qualitative interpretation of most solar observations, can be found in the book of Mitchell and Zemansky (1934) : a classical oscillator damped by its lifetime τ is submitted to the Larmor precession around the magnetic field direction with the angular velocity $\omega = eB/2m$; for the observer, there is a modification (generally a decrease) of the polarization degree and a rotation of the polarization direction, with respect to the field-free case, when $\omega\tau$ is of the order of 1 (see figure 3 in Tandberg-Hanssen, 1974a).

The quantum mechanical treatment is necessary for computing the polarization of scattered light for any spectral line and magnetic configuration. The Hanle effect appears as an interference phenomenon resulting from the overlapping of Zeeman sublevels. One finds again the condition $\omega\tau \approx 1$, which defines a typical value $B_{typ} = 2m/e\tau$ by comparing the Zeeman splitting with the natural width of the excited level (Sahal-Bréchet, 1981).

In the recent years major progress has been achieved in this field of theoretical research : House (1970) ; Sahal-Bréchet et al. (1977a, 1977b) ; Bommier et al. (1978, 1979a, 1979b, 1980, 1981, 1982) ; Stenflo (1977, 1978a, 1982) ; Landi degl'Innocenti (1982, 1983). On the other hand, useful observations had already been performed 50 years ago by Lyot (1934) and numerous data have been gathered in the last 15 years (see references in Leroy et al., 1977) owing to the availability of modern sophisticated polarimeters (Baur et al., 1981). It is possible now to have a general view on the capability of the Hanle effect as a method for measuring vector magnetic field (Stenflo, 1978b ; Sahal-Bréchet, 1981 ; Rees, 1982). But, before entering upon a more detailed discussion, it is worth remembering a very important feature which results from the very nature of the Hanle effect : one can expect useful magnetic field determinations via the Hanle effect only for media which are excited in a significantly anisotropic way (e.g. by a light or a particle beam). It is indeed a serious limitation which has to be considered for the selection of interesting astrophysical targets.

II FOUR INTERESTING FEATURES OF THE HANLE EFFECT

- 1 - The sensitivity is not systematically low when measuring small magnetic fields ; the range of good accuracy depends essentially on the spectral line under study.

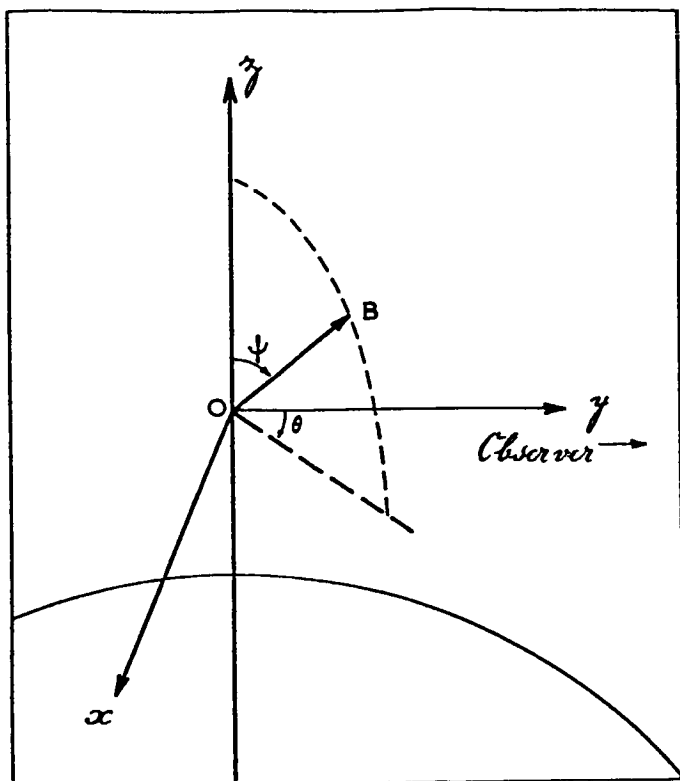
This feature, which is very different from the situation met in Zeeman measurements, is a consequence of the definition which has been given for B_{typ} : observing the Stokes parameters in a line with a B_{typ} of the same order as the field B to be measured gives the best conditions for the appearance of the Hanle effect. The following table, after Sahal-Br  chot (1981), gives B_{typ} for several lines of astrophysical interest : for permitted lines, τ is of the order of 10^{-8} s so that $B_{\text{typ}} \approx 10$ Gauss. But, for forbidden lines, τ is much longer and B_{typ} reaches very small values down to 10^{-5} Gauss.

If B is not too far from B_{typ} one can expect a good vector field diagnosis within the approximate range $0.1 - 10 B_{\text{typ}}$. If B is much larger than B_{typ} (as it is usually the case with forbidden lines) one cannot recover any information about the field strength but rather one gets the direction of the transverse component of the field. This is the so-called "strong-field case" which will be considered later.

line			B_{typ} (Gauss)	P_{lim}
Fe XIV	5303	$3p \ ^2P_{3/2} \longrightarrow 3p \ ^2P_{1/2}$	$5 \cdot 10^{-6}$	0.43
C III	1909	$2s2p \ ^3P_1 \longrightarrow 2s^2 \ ^1S_0$	$1.1 \cdot 10^{-5}$	1
He I	10830	$2p \ ^3P_{2,1,0} \longrightarrow 2s \ ^3S_1$	0.83	
He I (D_3)	5875	$3d^3 \ D_{3,2,1} \longrightarrow 2p \ ^3P_{2,1,0}$	6	
"	, major comp.	$3d \ ^3D_{3,2,1} \longrightarrow 2p \ ^3P_{2,1}$	6	
"	, minor comp.	$3d \ ^3D_1 \longrightarrow 2p \ ^3P_0$	16	1
C IV	1548	$2p \ ^2P_{3/2} \longrightarrow 2s \ ^2S_{1/2}$	22.5	0.43
N V	1239	$2p \ ^2P_{3/2} \longrightarrow 2s \ ^2S_{1/2}$	28.7	0.43
O VI	1032	$2p \ ^2P_{3/2} \longrightarrow 2s \ ^2S_{1/2}$	34.7	0.43
Si IV	1394	$3p \ ^2P_{3/2} \longrightarrow 3s \ ^2S_{1/2}$	78.2	0.43
Si III	1206	$3s3p \ ^1P_1 \longrightarrow 3s^2 \ ^1S_0$	295.	1
Ly $_{\alpha}$	1216		53.2	0.29
Ly $_{\beta}$	1026	, H $_{\alpha}$ 6563	16	
Ly $_{\gamma}$	992	, H $_{\beta}$ 4861	7	

- 2 - There is no difference in sensitivity for longitudinal and transverse fields relative to the observer. But there is a lack of sensitivity for magnetic directions parallel to the exciting beam.

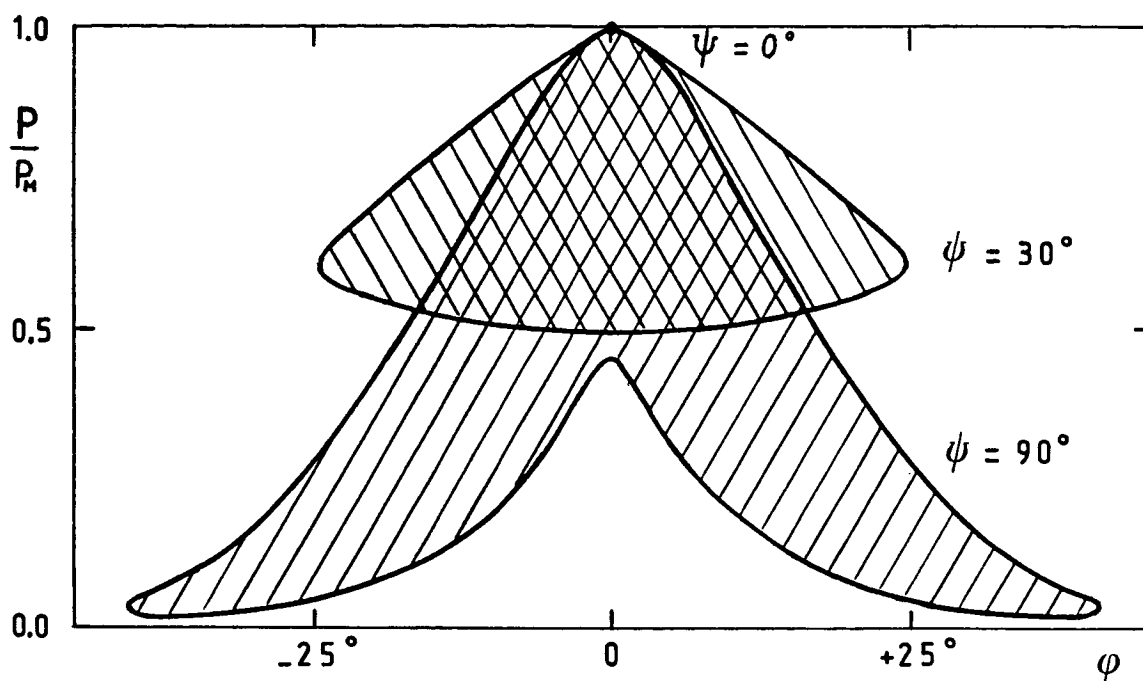
Throughout this paper we make use of the following reference system : the exciting beam travels along the z axis (it is the perpendicular to the photosphere in the case of the outer layers of the Sun). The observer is in the y direction. The field vector is characterized by the strength B , the angle ψ with respect to the z axis and the angle θ with respect to the line of sight, in the xy plane.



If $B = 0$ one observes the resonance polarization with a polarization direction perpendicular to the z axis (for permitted transitions) and a polarization degree P_{\max} . For increasing values of B one finds a polarization degree $P < P_{\max}$ and a rotation φ of the polarization direction. However, the Hanle effect does not show up if the field vector is parallel to the z axis : such a magnetic field, of course, cannot modify the symmetry which has been imposed by the exciting beam.

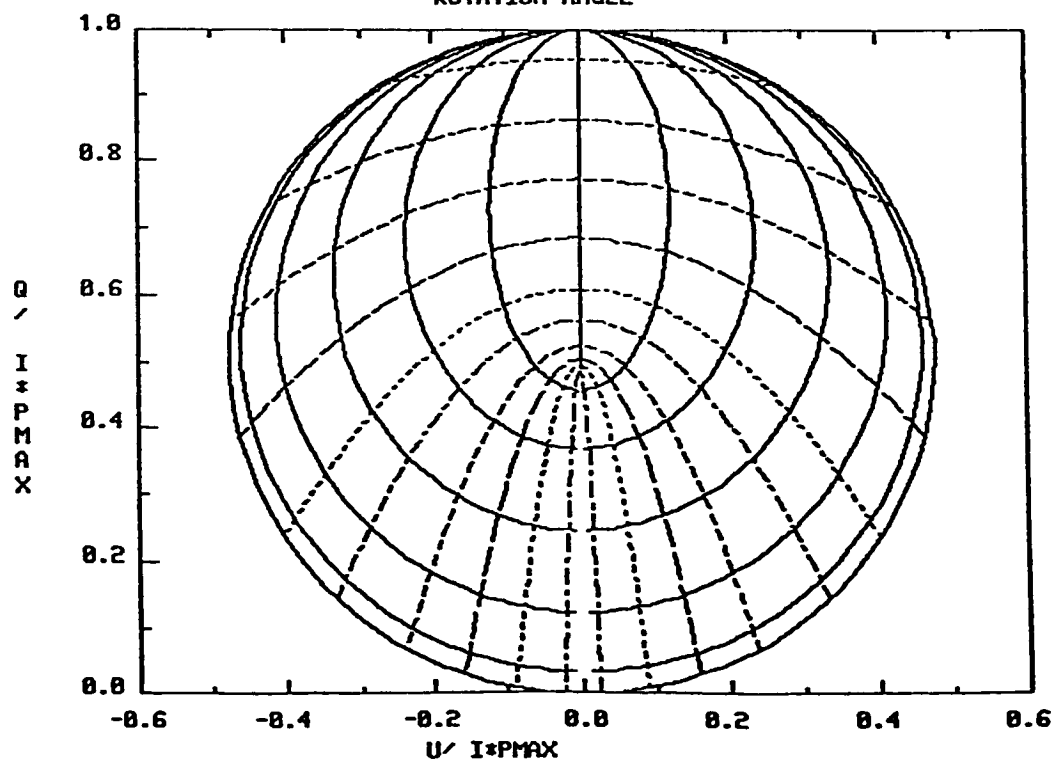
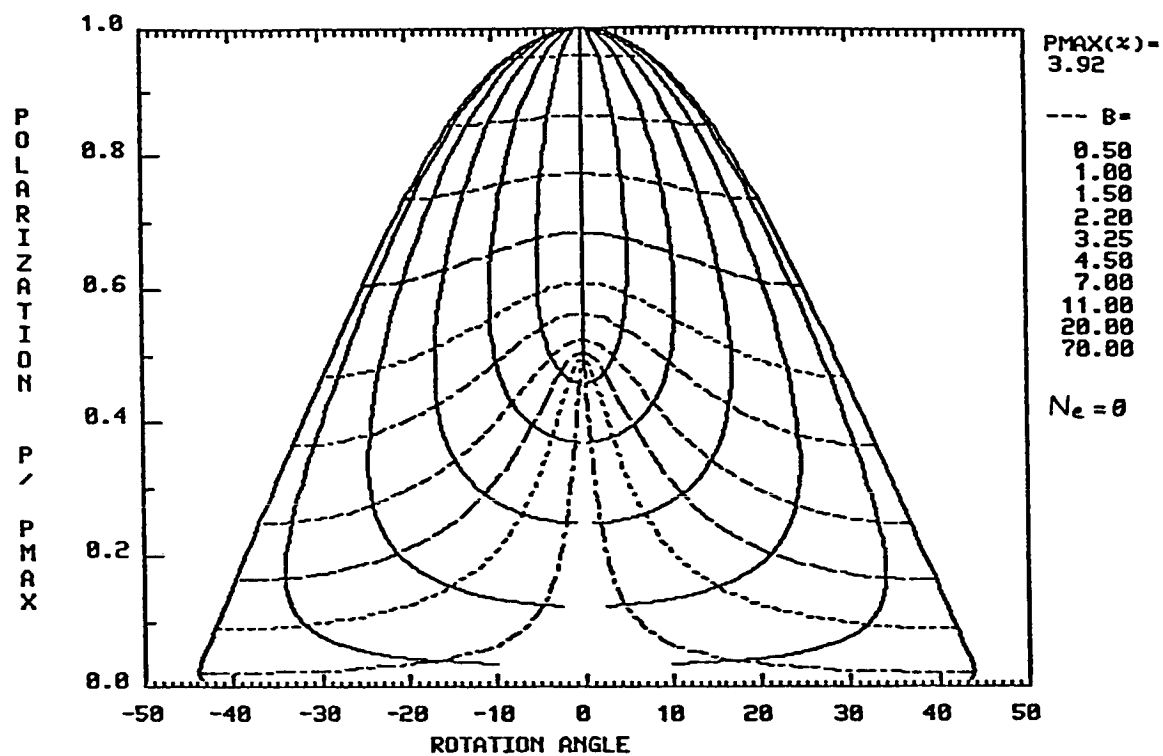
It is generally convenient to follow the phenomena on a "polarization diagram" (see on the next page) with P/P_{\max} in ordinate and φ in abscissa. When the field is parallel to the z axis, $\psi = 0$, one finds $P/P_{\max} = 1$ and $\varphi = 0$ for all possible values of B and θ . For increasing values of ψ one observes various values of P/P_{\max} and φ which depend on B and θ .

The parts of the polarization diagram relevant to the cases $\psi = 30^\circ$ and $\psi = 90^\circ$ are shown by shaded areas on the next figure. Obviously the change of the polarization parameters is more conspicuous when the field vector is roughly at right angle from the exciting beam ($\psi = 90^\circ$).



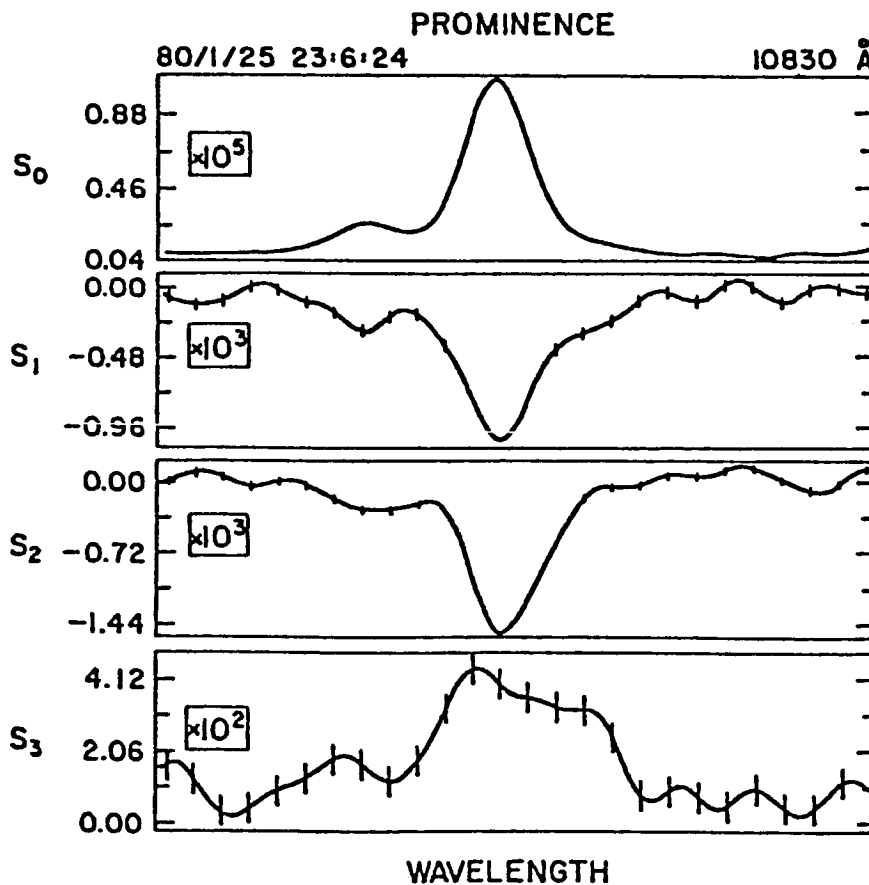
A more detailed polarization diagram computed for the $H\beta$ line (Bommier, in preparation) in the case $\psi = 90^\circ$ is displayed on the next page : dotted lines correspond to constant B values (specified in Gauss in the upper right corner) while full lines correspond to constant θ values (the outer curves refer to the cases of longitudinal fields, $\theta = 0$ on the right and $\theta = 180^\circ$ on the left ; the transverse field case, $\theta = 90^\circ$, yields the straight line with zero rotation angle). An alternative, useful, representation is given on the lower part of the figure where normalized values of the Stokes parameters Q and U are used along the ordinate and abscissa axis.

Summarizing the angular properties of the Hanle effect one can say that the magnetic field diagnosis can be very efficient whatever the field orientation is with respect to the observer, except when the field vector is close to the direction of the exciting beam. However, one must realize that the disappearance of the Hanle effect in this peculiar configuration ($\psi = 0$) still contains some information which may be useful : the fact that magnetic depolarization is observed in prominence spectra observed at the solar limb implies that the magnetic lines of force cannot be vertical !



3 - In the scattered light, the Stokes parameter V is generally one order of magnitude smaller than Q and U : observations refer essentially to the measurement of linear polarization.

This property of the Hanle effect is well visible on the following figure, after Baur et al. (1981) ; S_0 , S_1 , S_2 , S_3 are the four outputs of the "Stokes II" polarimeter which correspond to the Stokes parameters I, Q, U, V. This behaviour is not really surprising since the Hanle effect is nothing else than the progressive modification, for increasing field strength, of resonance scattering which is essentially a phenomenon yielding linear polarization. The origin of the small amount of circular polarization has been discussed by Landi degl'Innocenti (1982) ; anyway, for the magnetic field diagnosis we have to rely only on the observed values of Q and U, which has important instrumental consequences to be mentioned later.



- 4 - Various points in the line profile are affected in the same way by the change of linear polarization ; therefore, the polarization parameters can be measured on the integrated line profile.

This feature is obvious on the preceding figure. It has significant consequences for the instrumentation since it means that high spectral resolution is not needed for magnetic field measurements via the Hanle effect.

Further one must realize that we are left with only two observable quantities Q and U (since there is not any interesting wavelength variation) ; this is not sufficient to derive the three parameters B , ψ and θ which define the field vector. Therefore, unless some partial information on the magnetic field is already available, it is necessary to observe two different spectral lines in order to reach a complete knowledge of the magnetic field. Remember however that measurements via the Hanle effect in one line only are already more fruitful than the usual $B//$ Zeeman measurements.

III PRACTICAL MEASUREMENTS AND INTERPRETATION

1 - Useful spectral lines

The basic requirement has been found to measure the linear polarization of two different spectral lines. The reasons why selecting one line or another have already been discussed (Leroy et al., 1977 ; Sahal-Br  chot et al., 1977a ; Landi degl'Innocenti, 1982). One can summarize the favourable factors as following :

- The Zeeman structure of the line makes possible a scattering polarization (Beckers, 1974). The values of the maximum polarization degree P_{lim} for some interesting lines have already been given earlier in the table from Sahal-Br  chot (1981).
- The line emission is dominated by non-isotropic excitation processes which can be computed safely.
- The line intensity is sufficient to allow accurate measurements.
- The optical depth, however, remains small. We consider later what happens for large optical thicknesses.

In the solar prominences spectra the neutral helium lines were most favourable. Observing separately the two main components of the D3 λ 5875 A line has proved to be a good approach. The observation of both D3 and one Balmer line required a more sophisticated interpretation but has provided additional results (section IV).

2 - The accuracy of measurements

When the two lines to be observed are well apart in the spectrum it is convenient to perform the polarimetric analysis before some interference or birefringent filter which ensures a very good instrumental transmission. High spectral resolution asking for a spectrograph is wanted only in special circumstances (analysis of the two close components of the D3 line).

In the case of solar prominences the normalized integrated parameters Q/I and U/I were determined with some uncertainty mainly due to random errors $\Delta Q/I \approx \Delta U/I \approx 10^{-3}$ (Ratier, 1975 ; Baur et al., 1980 ; Baur et al., 1981). A comparison of the observations performed with one instrument at Pic du Midi Observatory and with another one at Sacramento Peak (Smartt, 1980) has shown a fairly good agreement, thus giving confidence to this type of measurements. It means also that the problem of instrumental polarization has been overcome in both instruments-which is not surprising for on-axis coronagraphs. It is important to stress that the linear polarization measurements required by the Hanle effect method practically exclude all those instruments including tilted mirrors before the analyzer (unless specially designed, accurate compensators are used ; see particularly Makita, 1970 and Makita et al., 1982).

3 - Final uncertainty on the vector field parameters

Having in one hand the computed polarization diagrams for the lines under study, and in the other hand the polarization measurements, it is possible to design an inversion procedure which yields the field vector and to estimate the final uncertainty of this determination. Such an analysis has been achieved first by Bommier et al. (1981) and then in a slightly different way by Landi degl'Innocenti (1982). From these investigations it follows that the prominence field orientation can be determined with an accuracy of some degrees ; remember that in the same situation ($B \approx 10$ Gauss) the Zeeman effect allows to measure only $B_{||}$ which tells almost nothing about the field orientation.

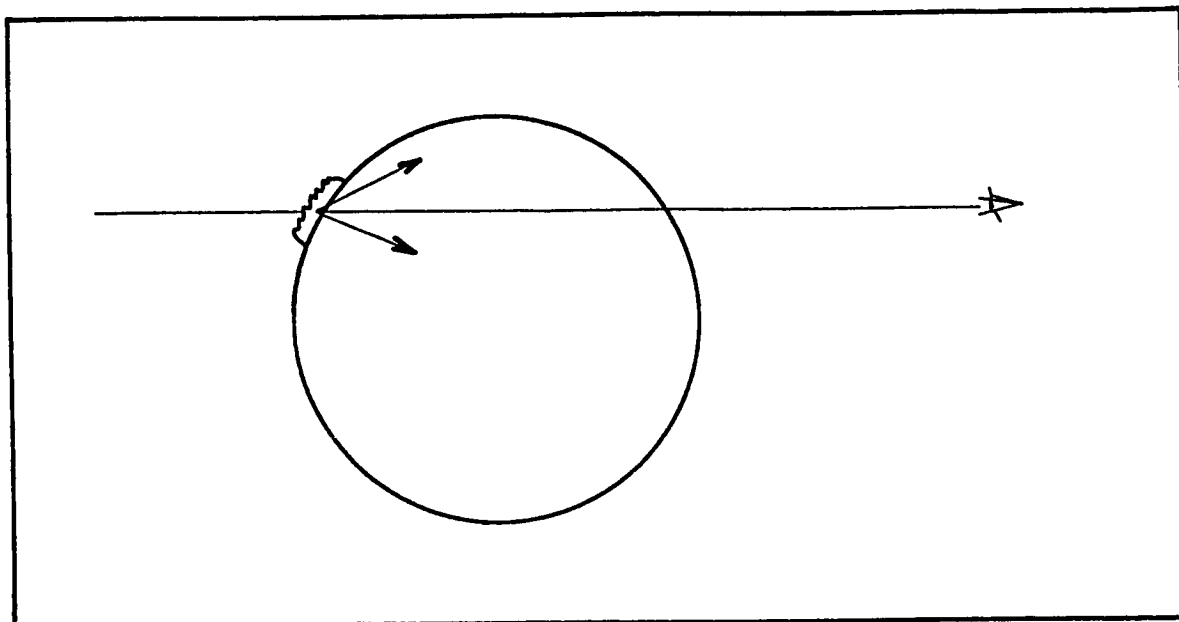
The reduction of measurements with instrumental noise has confirmed the predicted estimates of the uncertainty while pointing out some interesting differences between the three parameters B , ψ and θ to be determined. Roughly speaking one finds :

- ΔB about 0.5 Gauss (for $B \approx 10$ Gauss)
- $\Delta \theta$ of the order of 5°
- $\Delta \psi$ about 10°

In some circumstances noisy measurements can yield fictitious solutions for ψ but, most important, the results for B and θ remain essentially valid even for poorly determined ψ values. This peculiar behaviour obviously is linked to the general symmetry of the Hanle effect which has been described earlier.

4 - The 180° ambiguity on the transverse component of the field

Although this is generally not stated clearly, the Zeeman analysis of a line yields two possible solutions since there is always a 180° ambiguity on the transverse component of the field. This double solution occurs also for the Hanle effect and we are accustomed to remind that two magnetic vectors symmetrical about the line of sight have the same polarimetric signature (see on the next figure the geometric situation for observations at the solar limb).



Sahal-Bréchet (1981) has emphasized that this is a fundamental limitation (resulting, in fine, from the fact that the interaction matter-radiation is insensitive to the sense of propagation of the light) which can be avoided only if the scattering geometry can be modified in some fashion (Bommier and Sahal-Bréchet, 1979a). We come back soon to this question but I want to comment upon the solution which has been given to this problem in many solar investigations : in most cases the field solution close to the potential field seems to have been selected which may be a reasonable choice but should not be presented as resulting from an actual measurement ...

IV MORE COMPLICATED CASES : DRAWBACKS AND BENEFITS

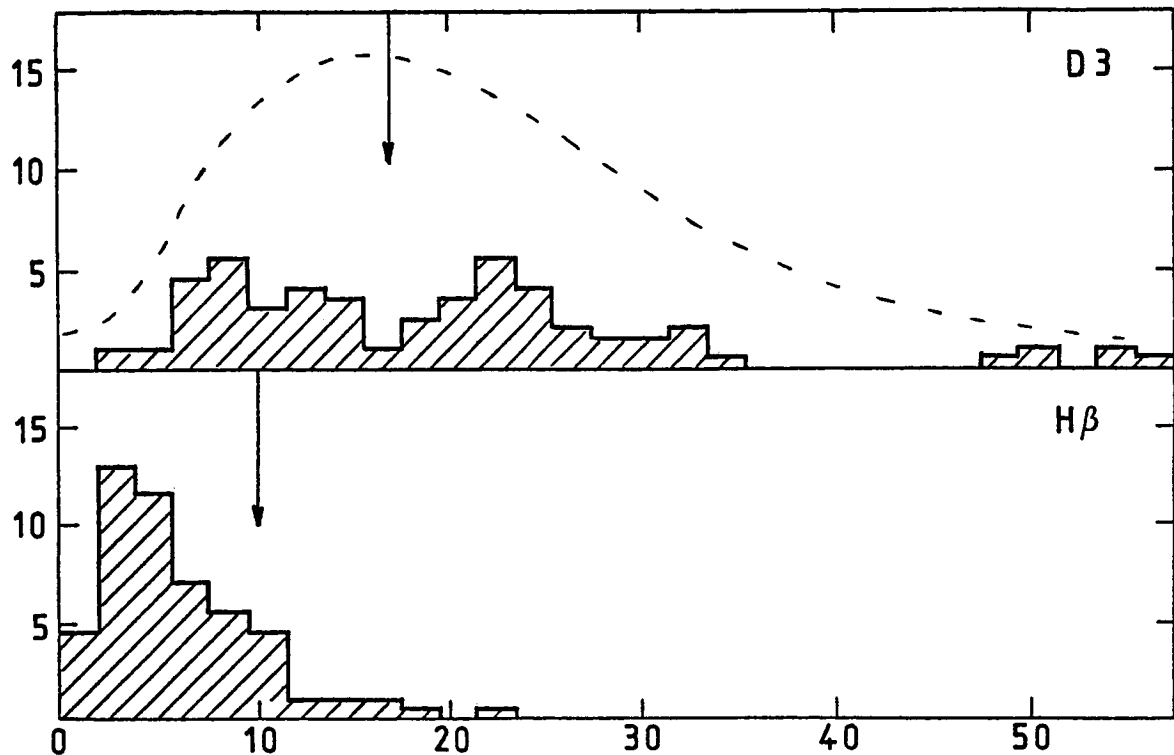
1 - The influence of collisions ; determination of N_e

The analysis of helium line formation in prominences has shown (Sahal-Bréchet et al., 1977a) that the collisions were negligible. This is no longer true for the hydrogen Balmer lines and a detailed theoretical analysis (Bommier, in preparation) has proven that there is an important non-magnetic depolarization of the $H\beta$ line which depends on the density in prominences. The observations made at Pic du Midi have revealed the same phenomenon as can be seen on the following figure.

$B = 8 \text{ Gauss}$

$\theta = 30^\circ$

$\psi = 90^\circ$



On these histograms the ordinates are the numbers of observed cases and the abscissas are the observed values of the polarization degree (in units of 10^{-3}). In the upper part, which refers to the helium D3 line, the dotted curve schematically recalls the distribution which has been observed for a very large number of prominences (Leroy et al., 1977); the shaded areas refer to a separate subset of observations which have been performed also for the H β line (lower histogram). The vertical arrow marks the polarization degree expected for average field parameters characteristic of quiescent prominences. It agrees with the upper histogram (which is obvious since the magnetic vector has been derived from the Hanle effect analysis of helium lines) but, on the lower histogram, the arrow forecasts a polarization degree three times larger than the observed one. Thus, it is clear that some depolarization mechanism, different from the Hanle effect, is acting on the H β emission. According to the theory of the Hanle effect for hydrogen lines, this effect must be assessed to depolarizing collisions.

Generally speaking there is no doubt that the measurements of polarization both in D3 and in H β lines yield four observed quantities, which is enough to derive the three components of the field vector and the electron density (which controls the collisions rate). The analysis of observations with the help of this newly developed theory (Bommier et al., in preparation) has revealed that the electron density in medium and high prominences is certainly around, and sometimes below, 10^{10} e/cm³, which is somewhat below the generally accepted figure. Although this particular analysis may appear to be disconnected from the vector field determination problem, it is worth remembering that the knowledge of both the magnetic field and the density is a compulsory starting point for any investigation on the magnetohydrostatic equilibrium of a plasma like solar prominences.

2 - The optically thick case

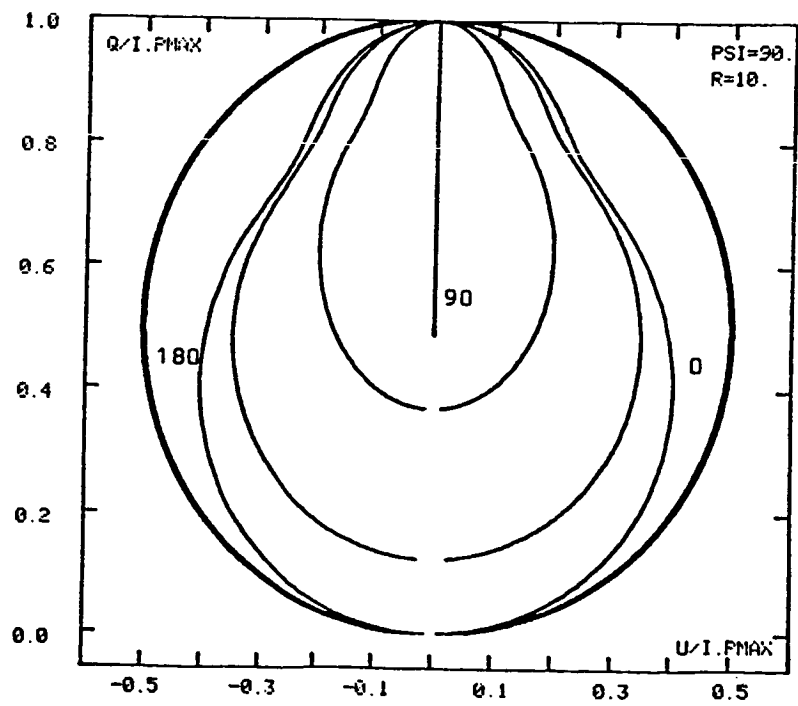
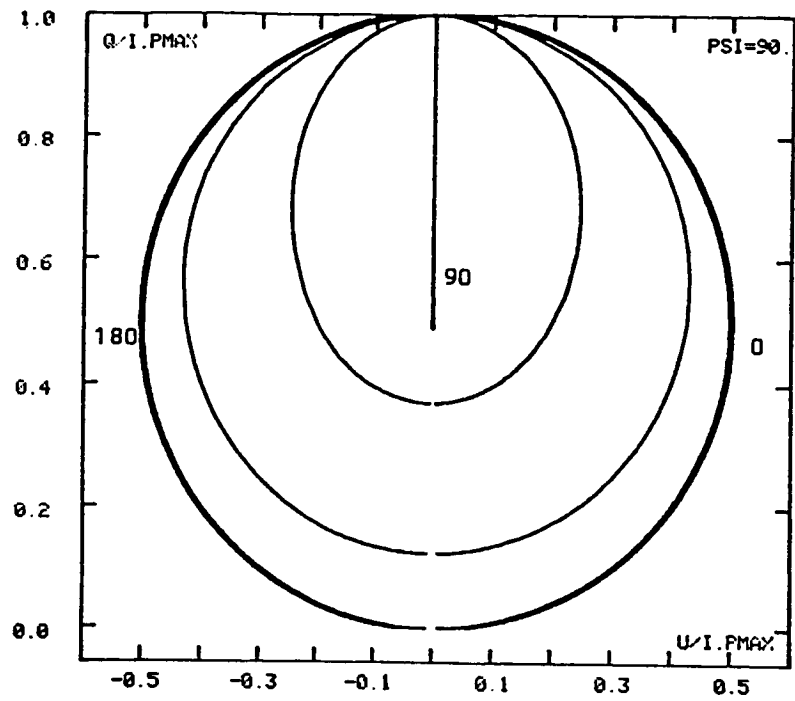
I do not want to dwell on this topic which is treated in V. Bommier's contribution. We have found at Pic du Midi (Leroy, 1981) that the polarization of prominences' H α line decreases for increasing optical depth which seemed, at first sight, consistent with the expected effect of multiple scattering.

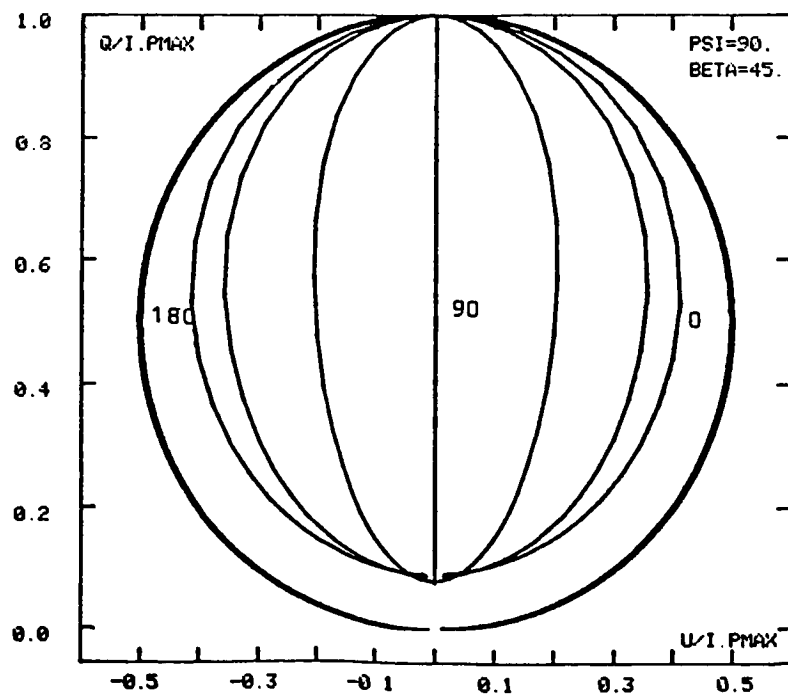
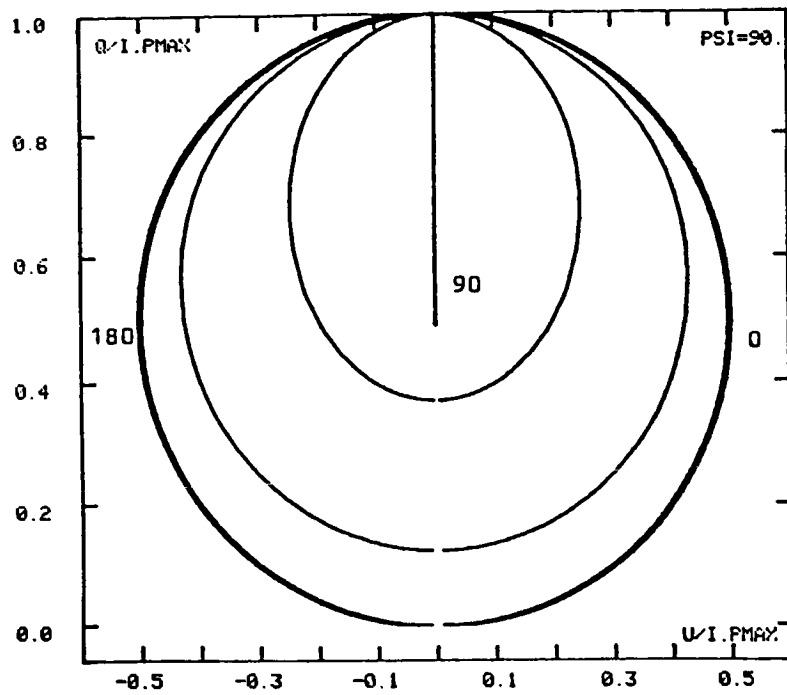
The Hanle effect for the H α line, in the non-optically thin case, has been investigated recently by Landi degl'Innocenti and Bommier (in preparation) and the general trend of the observations is well explained. Moreover, an interesting additional feature has arisen: in the optically thick case the scattering geometry is modified by the prominence's own emission and absorption; the degeneracy of solutions, which results in two undistinguishable solutions for the field vector, turns out to yield now two vectors which are no longer symmetrical about the line of sight and, eventually, it is possible to recover the field vector without any ambiguity. More details on this topics are to be found in the contribution by Bommier et al., in this Report.

3 - Magnetic fields with unresolved fine structure

Up to now it has been assumed that one is measuring the Hanle effect in homogeneous magnetic fields. As the fine structure of solar magnetic fields has been well established, at least at the photospheric level, it is important to investigate what can be deduced from Hanle effect analysis in the case of observations concerning fine-structure unresolved magnetic features. Starting from the formulas which give the Stokes parameters for a normal triplet (Bommier and Sahal-Br  chot, 1982) it is straightforward to compute the apparent polarization parameters which should be measured if some type of inhomogeneous magnetic structure is present.

On the next figure we give, in the upper part, the polarization diagram, for a normal triplet, in the case $\psi = 90^\circ$. Full lines have been drawn for several constant θ values, at 30° interval, from 0° to 180° . Increasing field strengths correspond to downwards shifts along the θ -constant curves.





Assuming now that the field strength is inhomogeneous, with the same total flux as previously, but with weak and strong values, in the ratio R , one obtains the lower diagram in the case $R = 10$. The most striking difference is that the outer limit of the polarization diagram is now closer to the central axis $U = 0$. One can conclude that large angular deviations of the polarization cannot result from magnetic fields with strong field strength variations.

In the next figure the upper diagram refers again to the homogeneous case. The lower diagram has been computed for a magnetic field with an unresolved helical structure (pitch angle $\beta = 45^\circ$) and a constant strength. This type of inhomogeneity yields a diagram different from the preceding case but, again, one can see that the outer boundary of the polarization diagram is moved inwards.

It is not difficult to see that this phenomenon must happen for any type of unresolved inhomogeneities, which, eventually, gives a possible diagnosis for homogeneous fields. Remember that the outer curves of the polarization diagram are obtained for longitudinal magnetic vectors ($\theta = 0$ or 180°): one can state that experimental points can be found on the outer boundary of the polarization diagram only if two conditions are fulfilled:

- a - the magnetic vector is nearly along the line of sight
- b - the magnetic structure is rather homogeneous

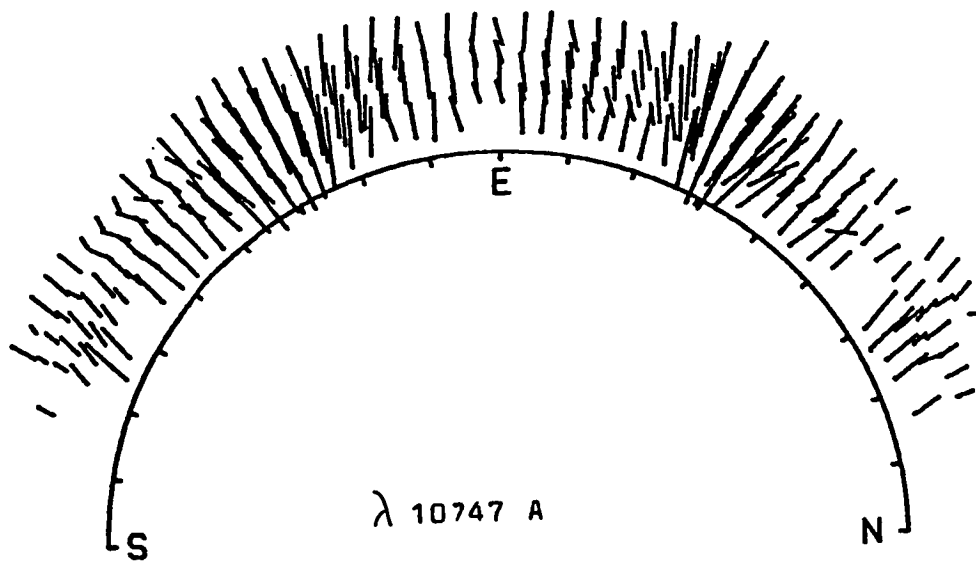
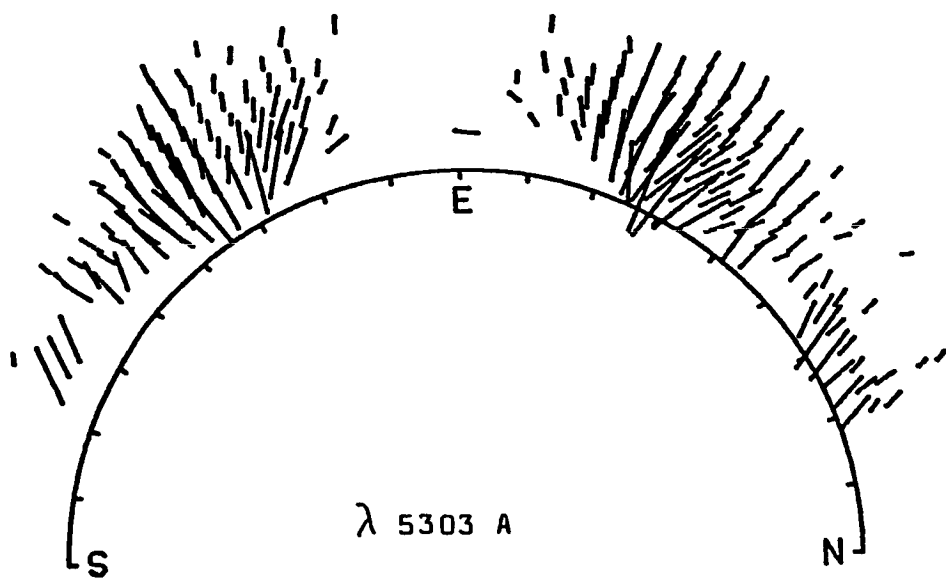
It was interesting to check this rule on our prominence observations: condition a can be fulfilled for high latitude prominences which are observed edge-on but condition b is generally considered to be unlikely in view of the highly structured appearance of prominence monochromatic pictures. However, we have found that 25 % of observations concerning edge-on prominences did give points on the outer edge of the polarization diagram: it means that conditions a and b were fulfilled for this important proportion of observations, a surprising result since the line of sight integration is maximum in such cases.

This is no place here to discuss solar prominences at length, but we must remember, in conclusion, that the Hanle analysis contains some useful information concerning the homogeneity of the magnetic field structure.

4 - The "strong-field case"

We have already mentioned that for $B \gg B_{\text{typ}}$ one cannot get any information on B via the Hanle effect except the direction of the field vector projected on the plane of the sky: it is parallel to the polarization direction for $\psi < 55^\circ$ and perpendicular for $\psi > 55^\circ$, a somewhat confusing behaviour for astrophysical applications.

A well known field of research relevant to the "strong-field case" is the solar corona: although B is not large (≈ 1 Gauss), B_{typ} is so small for the forbidden lines that the condition $B \gg B_{\text{typ}}$ is certainly fulfilled. Two instruments have been designed to measure the polarization of the Fe XIII (Sacramento Peak) and of the Fe XIV (Pic du Midi) lines. Their results are in very good agreement (see the next figure) which shows well the high accuracy of such methods. The interpretation, which is still in progress, is, however, complicated by line of sight integration effects (Querfeld, 1977; Arnaud, 1983).



As in the case of prominences, the study of the degree of polarization can give useful information on the local density because collisional excitation is not negligible in regard of radiative excitation. A recent analysis by Arnaud (1982, 1983) has shown that the measurements can be explained by a corona with rather high local densities, which imply a filling ratio as small as $1/5$ and a great deal of unresolved fine structure. On the other hand the conspicuous large scale structures which show up on the polarization maps during several consecutive days would be washed out if the magnetic field was changing drastically on small scale length. Therefore, as in the case of solar prominences, we are left with the apparent paradox of a strongly inhomogeneous density structure on one hand and a magnetic configuration dominated by large scale features on the other hand.

V ASTROPHYSICAL APPLICATIONS

1 - Prominences and corona

Prominences have been revealed to be ideally suited for the Hanle effect analysis and we have already given a lot of examples met in their study. More detailed conclusions are to be found in other papers by Bommier et al. Probably the most significant result has been (Leroy et al., 1984) to show that the field polarity across high prominences is opposite to the polarity observed at the photospheric level. Such a behaviour is consistent with the so-called Kuperus-Raadu model (Kuperus and Tandberg-Hanssen, 1967 ; Anzer, 1978) and we find that it fits well the observed features of prominences higher than 30000 km, and of all prominences located at high latitude (polar crown). However, the opposite magnetic structure, which would be consistent with the Kippenhahn-Schluter family of models (Tandberg-Hanssen, 1974b), is also met in a significant number of prominences which are found at low latitudes, among the objects lower than 30000 km.

Those findings give a first idea of the magnetic structure of the outer solar atmosphere above the photospheric neutral lines. The image which emerges is not simple and one wants, of course, to have further information on what happens in the hot coronal medium (e.g. what is the nature of the association between prominences and coronal helmets ? See for instance Pneuman, 1983).

2 - The turbulent photospheric field

Stenflo (1982) has observed the Hanle effect near the solar limb in the resonance λ 4227 Å line of calcium (a steep increase, in the very line center, of Q/I and U/I is the unambiguous signature of the Hanle effect ; see figure 6 in Stenflo's paper). Further, he has shown that the Hanle effect can be used to provide an estimate of the strength of unresolved turbulent magnetic field in the photosphere ; a first analysis of the depolarization observed near the limb has yielded a first order of magnitude result ($B_t \approx 10 - 100$ Gauss).

3 - Other solar observations

A very interesting result has been obtained by Henoux et al. (1983) : the conductive flux at the top of a solar flare can be derived from the observed polarization in the $H\alpha$ line and also in the $S\ I\ \lambda\ 1437\ \text{\AA}$ line (observed from SMM). Not only is this result important for the understanding of solar flares but it proves that polarization analysis of solar UV lines is now well possible with help of satellite instruments. It provides a promising way of determining the magnetic vector via the Hanle effect in the UV lines of the transition region (remember that the Zeeman method becomes insensitive in this wavelength region). Other applications can be forecasted for coronal researches ($Ly\alpha$ line for instance ; see Bommier and Sahal-Br  chot, 1982).

4 - Prospects for non-solar vector field determinations

Among the possible targets for Hanle effect analysis we have tentatively selected comets and planetary nebulae.

The role of resonance scattering in cometary spectra is well known. The study of the Hanle effect is certainly not straightforward since we are concerned with molecular spectra. However, the most troublesome circumstance is that the B_{typ} value for most bright cometary emissions is about 100 times larger than the expected B value to be measured (Crosley et al., 1968 ; Mies, 1975 ; Rahe, 1981).

In planetary nebulae we have a choice of large and small B_{typ} values with permitted and forbidden lines. The problem here is that little emission, if any, is expected to be due to resonance scattering, even though the flux of the central star has an essential role in the production of the nebular spectrum (Hummer and Seaton, 1973 ; Leibowitz, 1972).

At last, one can think about possible applications of the Hanle effect in stellar studies, particularly for stars with an outer shell (Be stars of course and possibly Mira type objects). As a matter of fact the Balmer lines polarization of such stars has already been measured (Mac Lean et al., 1979 ; Coyne and Mac Lean, 1979) but the interpretation has been restricted, up to now, to normal scattering processes in non spherical atmospheres. We look forward to undertake new observations which, however, are not so easy since it is necessary to make accurate spectropolarimetry with good spectral resolution on rather faint sources (see Landstreet and Angel, 1977).

At the present time it is not possible to say whether the Hanle effect analysis, which has been a powerful tool for studying the outer layers of the Sun, will be successful for investigations in stellar or nebular astronomy. However, there are not many available methods to derive the magnetic vector in stars and nebulae and we guess that it is worth trying, at least, some preliminary observations of this kind. First measurements are now in progress at Pic du Midi Observatory.

REFERENCES

- Anzer A., 1979, I.A.U. Colloquium n° 44, p 322
- Arnaud J., 1982a, Astron.Astrophys., 112, 350
- Arnaud J., 1982b, Astron.Astrophys., 116, 248
- Arnaud J., 1983, Thèse de Doctorat d'Etat, Université de Paris VI
- Baur T.G., House L.L., Hull H.K., 1980, Solar Physics, 65, 111
- Baur T.G., Elmore D.E., Lee R.H., Querfeld C.W., Rogers S.R., 1981, Solar Physics, 70, 395
- Beckers J.M., 1974, Solar Physics, 37, 351
- Bommier V., Sahal-Bréchot S., 1978, Astron.Astrophys., 69, 57
- Bommier V., Sahal-Bréchot S., 1979a, I.A.U. Colloquium n° 44, p 87
- Bommier V., 1979b, I.A.U. Colloquium n° 44, p 93
- Bommier V., 1980, Astron.Astrophys., 87, 109
- Bommier V., Leroy J.L., Sahal-Bréchot S., 1981, Astron.Astrophys., 100, 231
- Bommier V., Sahal-Bréchot S., 1982, Solar Physics, 78, 157
- Coyne G.V., Mac Lean I.S., 1979, I.A.U. Colloquium n° 46, p 386
- Crosley D.R., Tango W.J., Zare R.N., 1968, Ap.J., 154, L153
- Henoux J.C., Heritschi D., Chambe G., Machado M., Woodgate B., Shine R., Beckers J., 1983, Astron.Astrophys., 119, 233
- House L.L., 1970, I.A.U. Symposium n° 43, p 130
- Hummer D.G., Seaton M.G., 1973, Mem.Soc.Roy.Sci.Liège, 6ème série, V, p 225
- Kuperus M., Tandberg-Hanssen E., 1967, Solar Physics, 2, 39
- Landi degl'Innocenti E., 1982, Solar Physics, 79, 291
- Landi degl'Innocenti E., 1983, Solar Physics, 85, 3
- Landstreet J.D., Angel J.R.P., 1977, Ap.J., 211, 825
- Leibowitz E.M., 1972, M.N.R.A.S., 157, 115
- Leroy J.L., Ratier G., Bommier V., 1977, Astron.Astrophys., 54, 811
- Leroy J.L., 1981, Solar Physics, 71, 285
- Leroy J.L., Bommier V., Sahal-Bréchot S., 1984, Astron.Astrophys., 131, 33
- Lyot B., 1934, C.R.Acad.Sci.Paris, 198, 249
- Mac Lean I.S., Coyne G.V., Frecker J.E., Serkowski K., 1979, Ap.J., 228, 802
- Makita M., 1970, Annals Tokyo Astr.Obs., 2nd series, XII, n° 2, p 139
- Makita M., Hamana S., Kawakami H., Nishi K., 1982, Annals Tokyo Astr. Obs., 2nd series, XIX, n° 1, p 24
- Mies F.H., 1975, Ap.J., 202, 823
- Mitchell A.C., Zemansky M.W., 1934 (1961), Resonance radiation and excited atoms, Cambridge Univ. Press
- Pneuman G.W., 1983, Solar Physics, 88, 219
- Querfeld C.W., 1977, Reports Obs. Lund, n° 12, p 109
- Rahe J., 1981, Comets, in Landolt Bornstein, New series, VI, 2a, p 218
- Ratier G., 1975, Nouv.Revue Optique, 6, 149
- Rees D.E., 1982, Proc. A.S.A., 4, 335
- Sahal-Bréchot S., Bommier V., Leroy J.L., 1977a, Astron.Astrophys., 59, 223

Sahal-Bréchet S., Bommier V., 1977b, Reports Obs. Lund, n° 12, p 5
Sahal-Bréchet S., 1981, Space Sci.Rev., 30, 99
Smartt R.N., 1980, Private communication
Stenflo J.O., 1977, Reports Obs. Lund, n° 12, p 65
Stenflo J.O., 1978a, Astron.Astrophys., 66, 241
Stenflo J.O., 1978b, Rep.Prog.Phys., 41, 865
Stenflo J.O., 1982, Solar Physics, 80, 209
Tandberg-Hanssen E., 1974a, I.A.U. Colloquium n° 23, p 730
Tandberg-Hanssen E., 1974b, Solar Prominences, D.Reidel Publ.Comp.,
Dordrecht, Holland

SOME DESIGN CONSIDERATIONS

FOR A

SATELLITE-BORNE MAGNETOGRAPH

David M. Rust

The Johns Hopkins University Applied Physics Laboratory
Johns Hopkins Road, Laurel, MD 20707

ABSTRACT

This paper reviews the design criteria for a compact magnetograph that can monitor solar magnetic fields from a free-flying satellite for 5 - 10 years. The signal-to-noise ratio that can be obtained with a 10-cm f/10 refractor operated with a Fabry-Perot filter and a solid-state detector array is derived. With the candidate telescope, the longitudinal component of the magnetic field can be measured for the entire solar disk in a few minutes at a 20-G threshold and at 3-arcsec resolution. The candidate Fabry-Perot filter has a lithium niobate etalon, which can be tuned electrically and operated at a fixed tilt angle in such a manner that it cancels the solar rotational Doppler shifts in the transmitted spectrum. Principles of operation of various types of polarization modulators are presented, with the conclusion that photoelastic modulators and liquid-crystal devices hold the most promise for use in a satellite-borne magnetograph.

1. Introduction

A magnetograph can function from the ground. Why, then, consider building a satellite-borne magnetograph? In answer, we can point to two fundamental trends in solar magnetic field research that require observations from space. First, it is clear that most solar surface fields are concentrated in tiny bundles below the resolution limit of ground-based observations. These fields must be studied with large space telescopes. Second, there are a number of subjects that can be addressed only by uninterrupted full-disk monitoring:

- global, long-term circulation of magnetic flux
- evolution of active region magnetic fields (accurate accounting of the appearance and disappearance of magnetic flux),
- magnetic flux emergence and its relationship to flares,
- influence of surface fields and field changes on the structure of the corona.

The Solar Optical Telescope will address the problems requiring very high resolution and, because it will operate from the shuttle bay, its magnetograph can be heavy and temperamental and still be made to perform well. For the studies requiring full-disk monitoring, a magnetograph mounted in a free-flier in sun-synchronous orbit and having a lifetime of 5 - 10 years is needed. The prime design goals for this instrument are durability, simplicity, and compactness. Several approaches to meeting these goals for the important components -- telescope,

detector, filter and polarization modulator -- are described in this paper. And although the particular application is to longitudinal (i.e., line of sight) magnetic field measurements, the same considerations apply to transverse field measurements.

2. Telescope and Detector

The minimum objectives of the solar monitoring magnetograph are to measure the longitudinal component of the photospheric fields with an approximate 20-G threshold and to produce a full-disk magnetogram at least once per hour. Active region scans and global velocity measurements (for oscillation studies) must be obtained at higher cadences. A spatial resolution of 2 - 3 arcsec will accomplish the scientific objectives without exceeding the probable limitations on data transmission rates.

In order to make the magnetograph as durable as possible, all mechanical devices ought to be avoided, including image scanners, rotating filter elements, moving polarization analysers and tilting plates. Therefore, two-dimensional solid-state charge-coupled devices with 800 x 800 or 1000 x 1000 pixels are the detectors of choice, since they can record the whole solar disk without mechanical scanning. If a resolution of 4 arcsec is allowed, then the more readily available arrays with about 500 x 500 pixels can be used.

The choice of physical dimensions in solid state detector arrays is as wide as the choice of colors on early Ford cars. They are all about 10-mm square, so the optical system should produce a 10-mm solar image. For this, a simple refractor with a focal length of 1 m seems the best choice, since mirror coatings lose their reflectivity with time.

To minimize the size and weight of the magnetograph, the smallest aperture consistent with the signal requirements should be used. Three-arcsec resolution in the optical band can be obtained with a 10-cm diffraction-limited lens, but one must consider whether it will collect enough light to produce full-disk magnetograms at the required cadence.

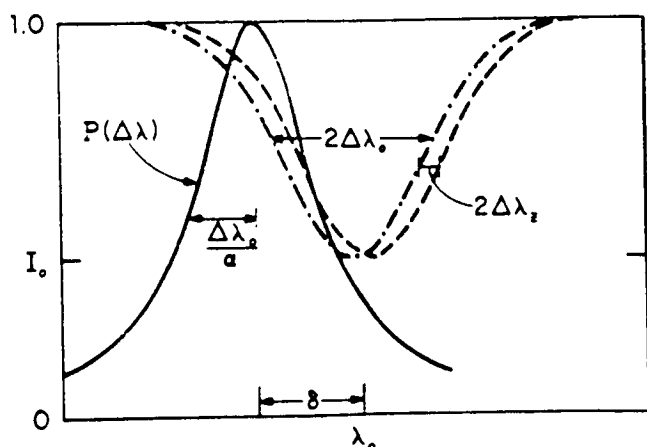


Figure 1. Profiles of the two circularly polarized components (dashed curves) of a magnetically sensitive spectral line and a filter profile (solid curve), $P(\lambda)$ that is offset by δ into the wings of the line. Optimum values of α and δ are evaluated in the text.

The efficiency of the optical system and detector can be estimated as follows:

<u>Element</u>	<u>Transmission(%)</u>
heat rejection filter	98
0.2 Å filter	25
objective lens	95
modulator and polarizer	30
blocking filter	25
camera lens	95
detector quantum efficiency	50

The net efficiency $\epsilon = 0.83\%$. If the 0.2 Å filter has a Lorentzian profile, which Fabry-Perot filters (see Section 3) do, then the signal-to-noise ratio can be found as follows, with reference to Figure 1: let f_λ be the continuum flux in $\text{erg cm}^{-2} \text{Å}^{-1} \text{s}^{-1}$ in the region of the observed line at wavelength λ . Then the photon flux at the telescope face is $5 \times 10^7 f_\lambda \lambda (\text{Å}) \text{ phot cm}^{-2} \text{s}^{-1}$. The Lorentzian profile of the filter passband is expressed as

$$P(\Delta\lambda) = \frac{1}{1 + \alpha^2 (\Delta\lambda / \Delta\lambda_0)^2} \quad (1)$$

where $\Delta\lambda$ is the wavelength displacement from the center of the passband, $\Delta\lambda_0$ is the HWHM of the line to be measured and α^{-1} is the filter HWHM in units of $\Delta\lambda_0$. Assume that the line profile is gaussian. The signal through the filter can be expressed as

$$S_0(\delta) = 2(1-I_0)\Delta\lambda_z \int_{-\infty}^{\infty} \frac{x e^{-x^2} dx}{1 + \alpha^2 (x - \delta/\Delta\lambda_0)^2}, \quad (2)$$

where $\Delta\lambda_z$ is the Zeeman shift and I_0 is the intensity at the center of the line in terms of the continuum intensity. The parameters δ and α shown here can be chosen as part of the instrument design. If we now let ϵ be the transmissivity of the optics times the quantum efficiency of the detector, and A the area of the telescope, then the net signal from the detector

$$S_T = 10^8 f_\lambda \lambda A \epsilon (1-I_0) \Delta\lambda_z \beta \text{ elect s}^{-1}, \quad (3)$$

where β is the integral above. β has been evaluated numerically for a range of values of α and $\delta/\Delta\lambda_0$. Peak values of β are obtained for $1 < \alpha < 4$ and for $\delta/\Delta\lambda_0 \approx 1$.

For 20-G fields, the magnetic signal $S = 6.8 \times 10^4 \text{ electrons s}^{-1}$ from each 3-arcsec pixel, while the photon flux produces a net charge of $\sim 8 \times 10^7 \text{ electrons s}^{-1}$ on each pixel. If the noise level is determined by photon statistics, the signal-to-noise ratio equals $\sim 7 \text{ t}^{1/2}$ at $\lambda 8648 \text{ Å}$. In principle, a good signal can be obtained in less than a second, but in practice, detector readout time will determine the cadence of the observations. The anticipated charging rate will saturate any of the solid-state detectors available today. About 100 readouts of a

device that saturates at 10^6 electrons will be required to collect the 8×10^7 electrons needed for a $S/N = 7$. With fast electronics, readout times for the whole sun (350,000 pixels) in the 1 s range can be achieved. Thus, full-disk magnetograms with at least a 20-G threshold can be obtained in a matter of a few minutes.

3. Filters

Since a spectrograph would be impractical in a satellite-borne magnetograph (too slow and susceptible to failure, because it requires that the image be scanned mechanically), one must consider ultra-narrow-band filters for spectral discrimination. Title (1984) has developed a universally tunable birefringent filter for use on shuttle missions, but birefringent filters will probably prove too large and too heavy to consider for a free-flying satellite. They have the added disadvantage of requiring mechanical devices for wavelength tuning.

Fabry-Perot etalons are lightweight and rugged and their spaceworthiness was established during the Skylab mission. However, the solid etalons used then did not have the flatness and uniformity required to yield a sufficiently narrow passband, i.e., one approximately equal to the halfwidth of the spectral lines used for magnetic field measurements. An air-spaced etalon, which is easier to fabricate, is being used for auroral studies with the Dynamics Explorer satellite, but for a number of reasons, a solid etalon is preferable for solar observations.

Figure 2 illustrates one advantage of solid etalons over air-spaced ones. The higher the refractive index of the spacer, the less the passband shifts for off-axis rays. The wavelength shift is given by

$$\Delta\lambda/\lambda = \phi^2/2n^2, \quad (4)$$

where ϕ is the angle between the incoming ray and the normal to the filter and n is the index of refraction of the spacer. This equation shows immediately that an etalon cannot be used with the $f/10$ telescope described in the previous section of this paper, since off-axis rays will broaden the filter profile to about 2 Å at 8468 Å, even with a lithium niobate (LiNbO_3) spacer ($n = 2.1$). A solution to this problem that will maintain high spectral purity for the entire disk is to place the filter in front of the objective, where maximum deviation of the rays from the normal is only 0.25 degrees. The passband of a Fabry-Perot filter with a LiNbO_3 spacer will shift only - 0.02 Å, then, for rays at the solar limbs, whereas it will shift by - 0.09 Å in an air-spaced etalon. The larger shift is unacceptable, since magnetically sensitive lines are < 0.2 Å wide.

Comparison of the off-axis behavior of a lithium niobate etalon and the Doppler shifts due to solar rotation shows that, if the etalon is operated at a tilt of 35 arcmin from the sun-center-to-telescope ray, the rotational Doppler shifts can be matched to $\Delta\lambda/\lambda \sim 2 \times 10^{-6}$, or 0.02 Å at 8468 Å. Thus, a solid etalon with a LiNbO_3 spacer can provide a passband that is narrow and correctly positioned on the wing of the line simultaneously everywhere on the visible disk. For research on Jupiter's atmosphere, Trauger and Roesler (1972) have used the tilted Fabry-Perot technique to remove line smearing due to the 12 km s^{-1} rotation rate of that planet.

There are difficulties with the before-the-telescope design that have yet to be resolved. Lithium niobate is a birefringent material, so analysis of the polarization must be accomplished before the light passes through it. And the filter (or the polarization analyser) will determine the aperture of the instrument. If 3-arcsec spatial resolution and adequate throughput are to be maintained,

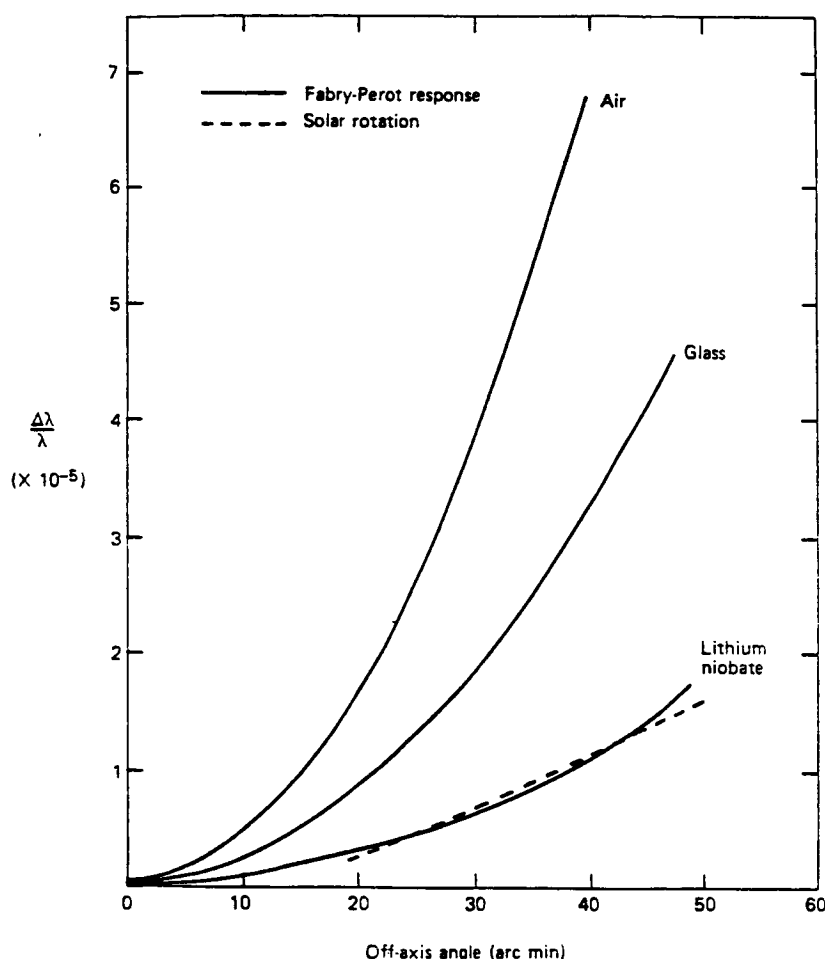


Figure 2. Shift $\Delta\lambda/\lambda$ of the transmission peak of a Fabry-Perot etalon vs. angle between an incident ray and the normal to the etalon face. Higher index spacers lessen the angular sensitivity. Solar rotation (dashed line) produces a $\Delta\lambda/\lambda$ across the visible disk that can be compensated for in a tilted lithium niobate Fabry-Perot.

then one must be able to fabricate a 10-cm diameter etalon. To give the required spectral resolution, the etalon must be flat and parallel to 1/50th wave, or better (Ramsay, 1969). To facilitate separation of orders with a thin-film filter, the etalon must be only 0.25 mm thick.

Fabrication of a LiNbO_3 etalon with the prescribed aspect ratio has not yet been attempted, but large-diameter glass etalons have been made by Harihan, Oreb and Leistner (1984), and Gunning (1982) has made a 10-mm double-cavity lithium niobate Fabry-Perot filter. In addition, Gunning demonstrated the principal advantage of LiNbO_3 etalons, i.e., electrooptic tunability.

Mechanical means are usually used to tune Fabry-Perot filters, either by varying the spacing between the mirrors, for air-spaced filters, or by tilting. In LiNbO_3 , application of an electric field along the optic axis induces a change in the refractive index of $\delta n = -1/2 n_0^3 r_{13} E_3$ for light propagating along the optic axis. n_0 is the refractive index in the absence of a field, r_{13} is the relevant electrooptic coefficient, and E_3 is the applied field. The voltage required to scan the filter passband over one free spectral range, the halfwave voltage, is

$$V_{1/2} = \pm \lambda/2 n_0^3 r_{13}. \quad (5)$$

The free spectral range of a 0.25-mm LiNbO_3 etalon is 6.8 Å at 8468 Å, and since $r_{13} = 8 \times 10^{-12} \text{ m V}^{-1}$, ~ 5300 V would be required for a fully tunable filter.

Fortunately, the tuning requirements of a solar magnetograph are modest. Only temperature drifts in the instrument and Doppler shifts due to orbital motion of the satellite must be compensated for and both are $< 0.2 \text{ Å}$. And, for photospheric velocity measurements, it is usually desirable to shift the passband from one wing of the spectral line to the other, and this also requires ~0.2 Å tunability. Thus, only $\pm 100 \text{ V}$ will provide adequate control over the passband. Higher voltages would be needed in a more versatile instrument, one that can be used in conjunction with several narrow-band order-isolation filters to study many spectral lines.

4. Polarization Modulators

New approaches to the task of detecting the polarization state of the incident solar light will be required if a satellite-borne magnetograph is to operate for very long. Hagyard, Cumings and West (1983) have described the problems encountered with the frequently-used potassium dideuterium phosphate (KDP) crystal modulators. These include hysteresis effects, breakdown under the high voltages required to achieve modulation, and a very narrow-angle acceptance cone.

Makita (1984) has described problems inherent in the use of rotating waveplates. They introduce a spurious modulation in the transmitted beam because reflection at their surfaces depends upon the orientation of the crystal axes relative to the polarization of the incident beam. And, mechanisms should be avoided for a long-lived instrument.

Probably the most indestructible polarization modulator is the fused quartz photoelastic modulator (PEM) perfected by Kemp (1969). Inducing a strain in a quartz plate will induce birefringence that can modify the polarization state of light propagating perpendicular to the strain direction. A drawback is that a PEM operates most efficiently at the resonant mechanical vibration frequency of the quartz plate, which is about 50 kHz for a plate several centimeters across. To reduce the modulation to a rate comparable to a practical readout rate, ~ 60 Hz, for two-dimensional array detectors, two PEM's whose resonant frequencies differ by the readout rate may be used. Alternating transmission, at the beat frequency, of right- (I_R) and left-circularly (I_L) polarized light by a pair of PEM's separated by a linear polarizer and a quarter-wave plate is illustrated in Figure 3. The figure shows the complete sequence of operations for a magnetograph with a charge-injection device (CID) detector which is "injected" (cleared) just as either the left- or right-circular signal is increasing. The signal is integrated for about 20 ms, then read out for 3 ms.

PEM's operate at only 150 volts and two of them consume only ~0.25 watts. They have very large acceptance angles and are completely free from the residual birefringence and hysteresis effects that plague KDP crystal modulators. Their use in a satellite-borne magnetograph should be investigated further.

Another low voltage modulator that is now commercially available is made of transparent lanthanum-modified lead zirconate titanate (PLZT), a ferroelectric ceramic that displays the Kerr effect (Haertling and Land, 1971; Cutchen *et al.*, 1975). Since the Kerr effect is a phase retardation proportional to the square of the applied electric field, a PLZT plate alone cannot give, say $\pm \lambda/4$ retardation. It must be operated with a fixed retardation waveplate, and therein lies a problem. A symmetric error in retardance, caused, for example, by a slightly incorrect

voltage, lowers a magnetograph's signal to noise ratio slightly, but it does not cause a false signal. In a quadratic device, any error in the applied voltage will create an artificial signal that must be removed later.

At APL, we have studied the retardance of a commercial PLZT unit and found hysteresis effects that remain for up to 30 s. Another problem is that metal electrodes, typically 76- μm thick lines, must be imbedded across the face of the plate. PLZT modulators, therefore, cannot be used near an image plane; elsewhere, they will reduce the system modulation transfer function. Lifetime tests (Cutchen *et al.*, 1975) show that the switching time of PLZT rises sharply after about 500,000 cycles, which is only ~ 3 h at 60 Hz!

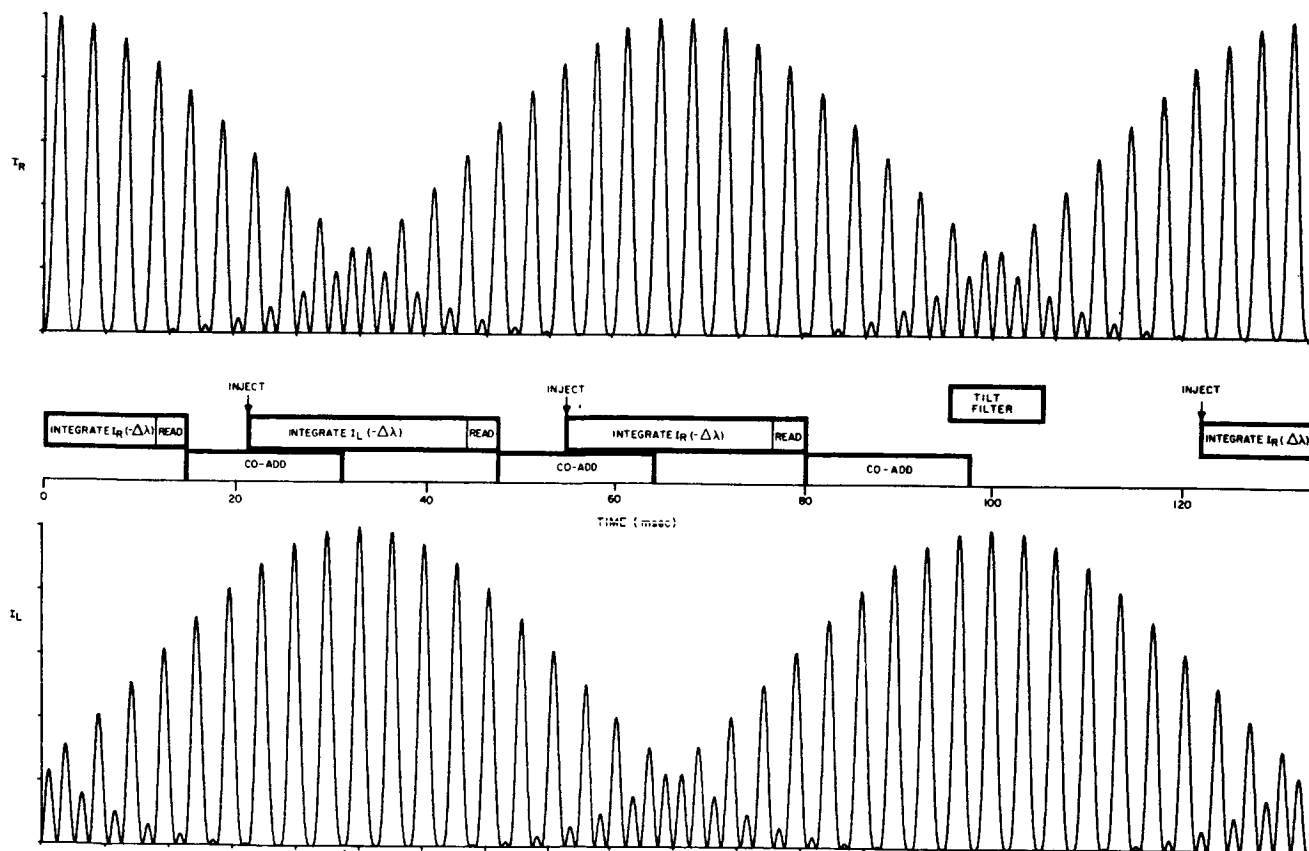


Figure 3. Right- and left-circularly polarized signals (I_R and I_L) passed by two photoelastic modulators separated by a fixed circular analyzer and followed by a polarizer. The time line in the center shows signal integration, readout, and computing periods.

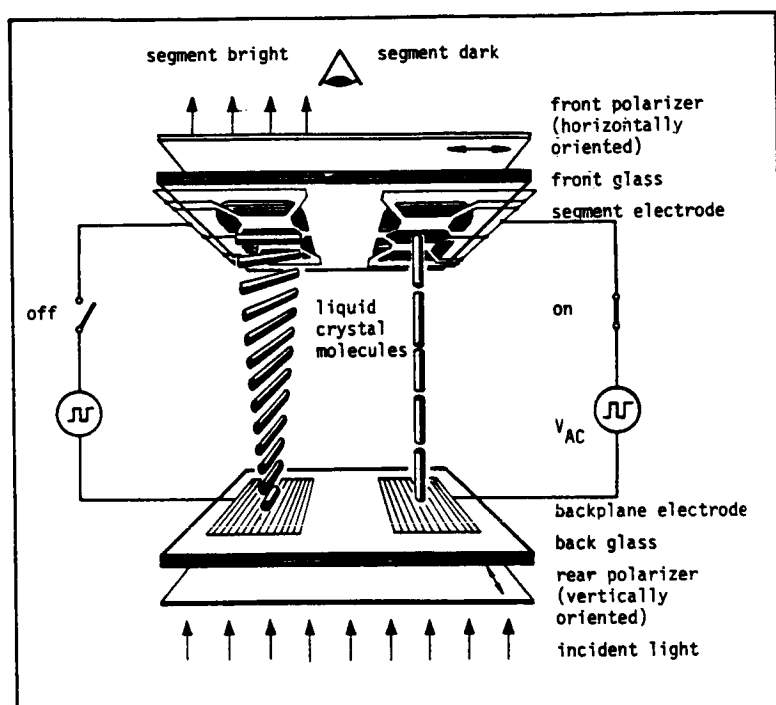


Figure 4. Operation of a twisted-nematic liquid-crystal display. Light entering the unenergized digit on the left has its plane of polarization rotated by 90° by the twisted-nematic liquid crystal layer and so passes through the crossed polarizers. Light entering the energized digit to the right is not affected by the realigned liquid-crystal molecules, so it is blocked by the exit polarizer. Thickness of the liquid-crystal layer is 6 - 10 μm (Courtesy Videlec Ltd).

Finally, we have been investigating twisted-nematic liquid-crystal devices. Typical LCD operation is sketched in Figure 4. In the absence of an electric field, light polarized in a direction parallel or perpendicular to the axes of a typical device is rotated by the crystal in a waveguide mode through 90° , i.e., the device resembles in its effect a half-wave plate whose axis is at 45° to the plane of polarization. However, as Figure 5 shows, when the plane of polarization of the incident beam is at an angle other than 0° or 90° to the device' axes, the retardation is a strong function of position in the device plane. The reason is that the indices of refraction for light polarized perpendicular to and parallel with the "director" of the molecules (represented as rods in Figure 4) are not the same. The retardance of the light, then, is strongly dependent upon the thickness of the liquid crystal layer, which is difficult to control. For a uniform effect, the incident light must be plane polarized along one of the device axes, as shown in Figure 4.

Modulation in an LCD is achieved by destroying the twisted crystal structure by application of an electric potential, typically 5 - 6 volts for a 10- μm layer. Our tests show that LCD's are remarkably insensitive to pressure, temperature, and minor voltage errors in their ability to modulate plane polarized light when operated as shown in Figure 4. More details about liquid crystal devices may be found in articles by Baur (1981), Clerc (1981) and Scheffer (1978).

Because LCD's are either passive elements (electric field applied) or waveguides (field off), they are achromatic. This is an important advantage in a magnetograph destined to operate at many lines in the spectrum. To achieve modulation of circularly polarized light, however, an LCD must be used with a fixed quarter-wave plate.

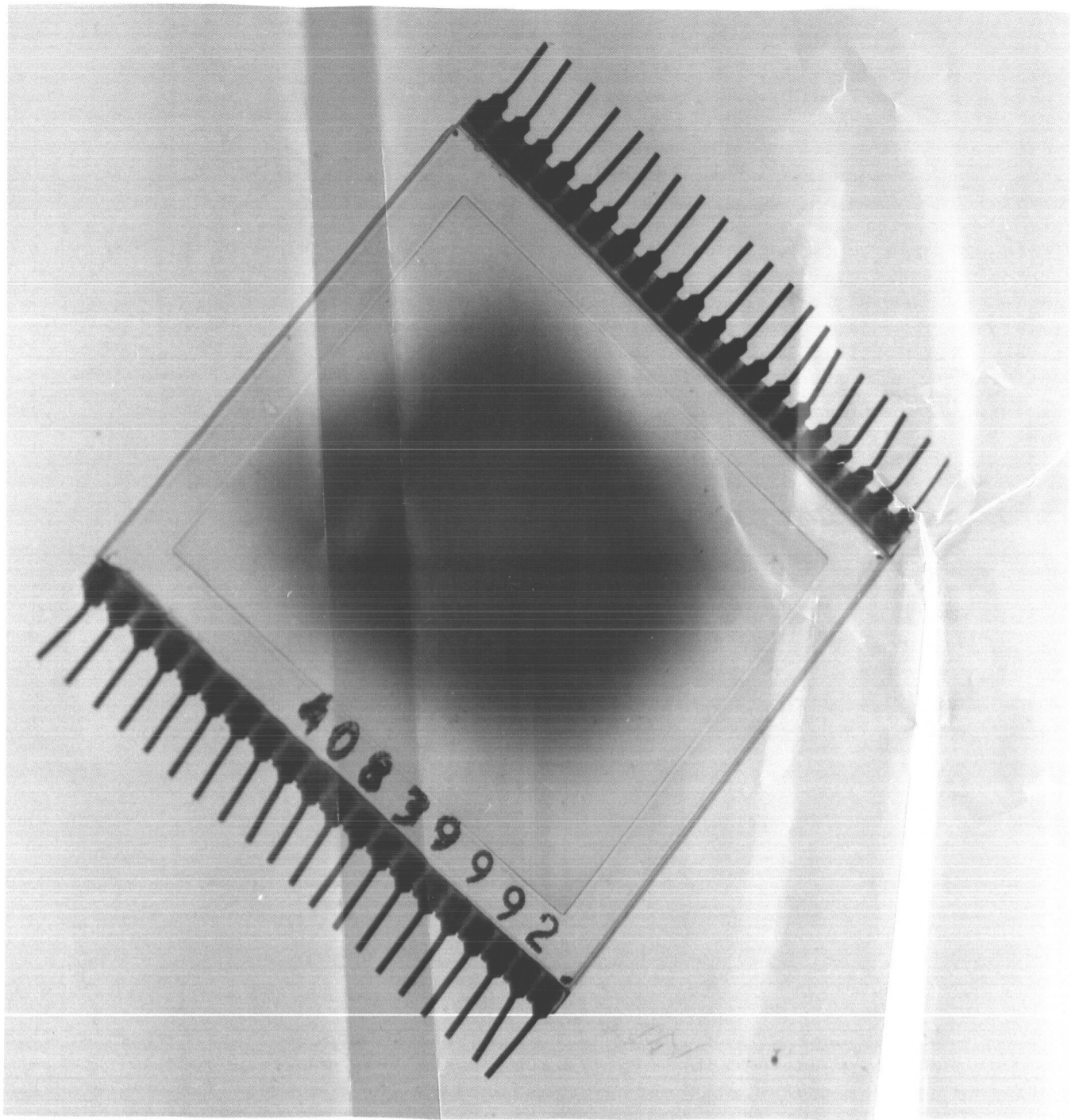
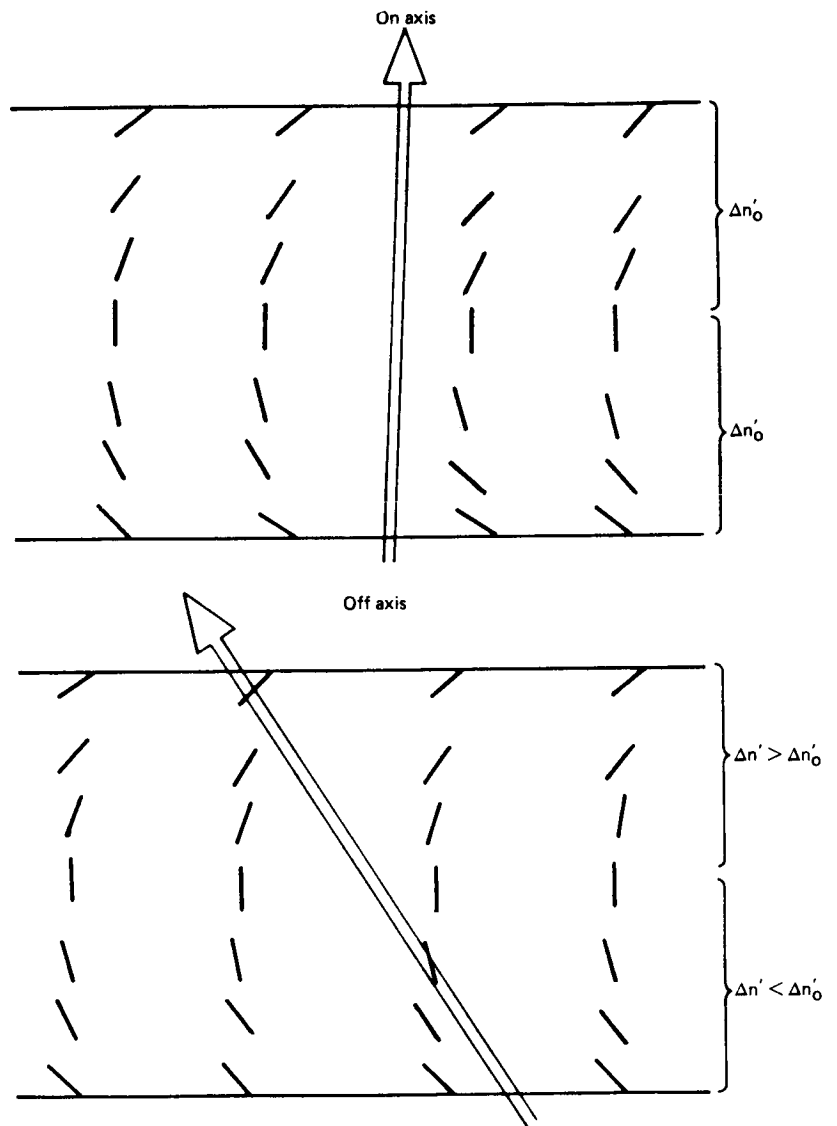


Figure 5. Interference colors in a commercial LCD placed between two parallel polarizers and set at a 45° angle to the plane of polarization. The colors show that the optical path difference ranges from about $0.3 \mu\text{m}$ (yellow) on the edges to $0.6 \mu\text{m}$ (blue) at the center. The same unit showed excellent uniformity, however, when operated at 0° and 90° to the plane of polarization.

Figure 6. This sketch illustrates how the wide-angle response of the Tektronix pi cell LCD is achieved. The obliquely incident ray will be more nearly parallel to the directors (short vectors) of the liquid-crystal molecules in the bottom half of the device than for normal incidence rays but, in the top half, the oblique ray will be at a greater angle to the directors. The net retardation, therefore, is relatively insensitive to beam angle. $\Delta n'$ is the effective birefringence of the liquid crystal.



The response time for most LCD's is 100 ms, which is rather slow for light modulation in a magnetograph. This drawback has been overcome recently with the introduction of a device with a 2-ms response time (Bos, Johnson and Koehler/Beran, 1983). The new device also has a very high acceptance angle, since, as Figure 6 shows, the surfaces that confine the liquid crystal have been designed to orient the molecular directors to give only a small net deviation from the on-axis index of refraction, regardless of ray angle. The contrast ratio between the "on" state and the "off" state transmission is $\sim 30:1$.

Tektronix claims that the new LCD has an expected lifetime of 10 years. Probably the principal factor limiting lifetime is breakdown of the complex liquid crystal molecules. Their behavior in the radiation fields encountered in space is unknown.

5. Conclusions

The new technologies for detectors, filters, and polarization modulators that have become available recently are important for realization of a small, lightweight and durable magnetograph for satellite-based operation. With them, 3-arcsec measurements of the longitudinal field component at a minimum detectable field of 20 G can be made over the whole solar disk in a few minutes with a 10-cm f/10 refractor and a 800 x 800 solid-state detector array. Spectral selection can be accomplished with a solid Fabry-Perot etalon located in front of the telescope objective. Lithium niobate may make an excellent etalon spacer, if optical quality material can be grown to a diameter of ~10 cm. Smaller diameter etalons are definitely feasible, but their use in a magnetograph would require that the image be scanned mechanically, a drawback in instruments intended for a 5 - 10 year lifetime.

The search for the perfect electrooptic polarization modulator continues, since none of the devices considered here, KDP's, quartz plates in oscillation, PLZT wafers, or LCD's operates reliably at low voltage, for long intervals, at any retardation, and with a wide acceptance angle. A recently introduced LCD holds the most promise, having only the drawback that it must be operated in conjunction with a fixed retardation plate. Since this LCD can be fabricated with a 10-cm or larger aperture, we plan to install it in a new magnetograph at APL (Figure 7) for operational tests.

ACKNOWLEDGEMENTS

This work was sponsored by National Aeronautics and Space Administration grant number NAGW 522. I am grateful to R. Smartt of the National Solar Observatory and J. McCormick of Tektronix, Inc. for the use of filters and a pi-cell LCD, respectively, and to R. Cohn for help in the experimental program.

REFERENCES

- Baur, G., 1981, Mol. Cryst. Liq. Cryst. **63**, 45.
Bos, P.J., Johnson, P.A., Jr., and Koehler/Beran, K.R., 1983, SID 83 Digest, **30**.
Clerc, F., 1981, Displays, Oct. issue, p. 341.
Cutchen, J.T., Harris, J.O., Jr., and Laguna, G.R., 1975, Appl. Optics **14**, 1866.
Gunning, W., 1982, Appl. Optics **21**, 3129.
Haertling, G.H. and Land, D.E., 1971, J.Amer. Ceramic Soc. **54**, 1.
Hagyard, M.J., Cumings, N.P., and West, E.A., 1984, Proc. Kunming Workshop on Solar Physics and Interplanetary Travelling Phenomena (in press).
Harihan, P., Oreb, b. F. and Leistner, A. J. 1984, Opt. Eng. **23**, 294.
Kemp, J. C. 1969, J. Opt. Soc. Amer **59**, 950.
Makita, M., 1984, Workshop on Measurements of Solar Vector Magnetic Fields, Marshall Space Flight Center, Huntsville, Al.
Ramsey, J. V. 1969, Appl. Optics **8**, 569.
Scheffer, T.J., 1978, Inst. Phys. Conf. Ser. No. 40, 65.
Trauger, J.T. and Roesler, F.L., 1972, Appl. Optics **11**, 1964.
Title, A. 1984, Adv. Space Res. -Proc. XXV COSPAR Mg., Graz. Austria (to be published).

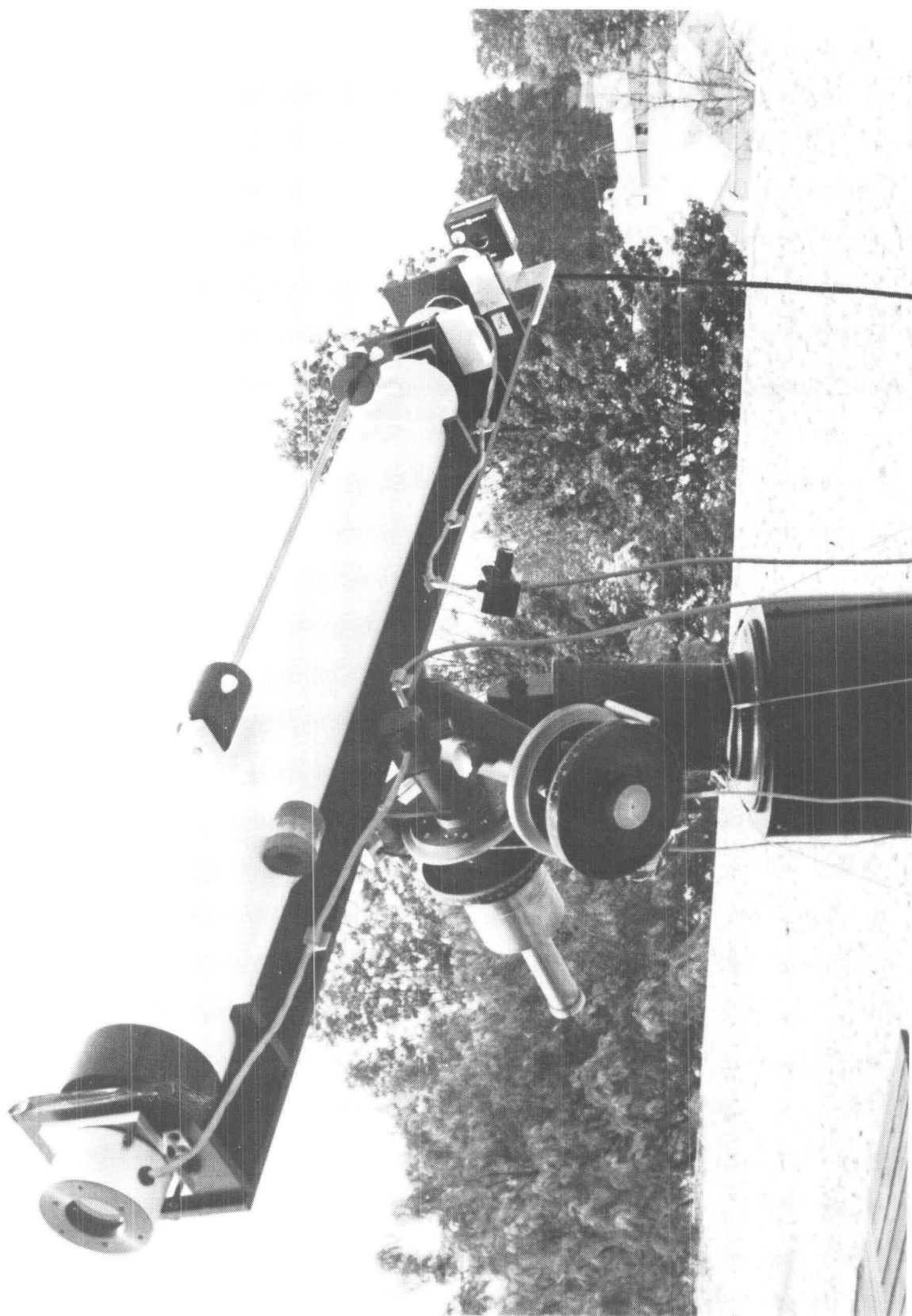


Figure 7. A 100-mm f/10 refractor on the roof of the Research Center at APL. Components to be tested on this spar-mounted magnetograph can be interchanged easily. A Fabry-Perot filter (left) is mounted in front of the objective lens. A CID video camera (right) relays a full disk solar image to a digitizer and image processor. The telescope was made by the Dobbins Instrument Company, Ohio.

MEASUREMENTS OF VECTOR FIELDS WITH A DIODE ARRAY

E. J. Wiehr

Universitäts-Sternwarte, D-3400 Göttingen

and

W. Scholiers

Institut d'Astrophysique, B-1050 Bruxelles

Apparatus

A polarimeter has been designed for high spatial and spectral resolution. It consists of a quarter-wave plate alternately operating in two positions for Stoke-V measurements; and, resp., an additional quarter-wave plate for Stokes-U and -Q measurements. At the spectrum plane a 100 x 100 pixel Reticon is used as receiver. The geometric scale of the spectrum was adapted to the dimension of the diode array by a minification optics yielding about 0.75 arcsec x 18 mÅ (at 6300 Å) on each 60μ x 60μ pixel. The spatial range thus covers 75 arcsec, the spectral window of about 1.8 Å allows the simultaneous observations of neighbouring lines. The block diagram of the data processing and acquisition system consists of five memories each one having a capacity of 10⁴ 16-bit

words. One is used to store the raw data, one the dark current (zero intensity signal), the third one stores the gain (flat field calibration), the two remaining memories are used to add a chosen number of successive measurements corresponding to two opposite states of polarization. The read-out is synchronized with the rotation of a $\lambda/4$ plate between two positions. Display equipment visualizes the corrected Stokes profiles in real time on a screen and allows the observer to store the best ones on a floppy disc. This equipment has been operated at the Locarno solar station of the Göttingen observatory (Wiehr et al., 1980).

The total time to acquire profiles of Stokes parameters can be chosen by selecting the number of successive measurements added in the memories, each individual measurement corresponding to an integration time of 0.5 sec. Typical values range between 2 and 60 sec depending on the brightness of the structure, the amount of polarization and a compromise between desired signal-to-noise ratio and spatial resolution.

Compensation of Telescopic Polarization Effects

The reflection on two plane mirrors of the Coudé telescope introduces instrumental polarization (Wiehr 1971), which can be described by a 16-term matrix. Evaluation of this matrix shows that the main effect is a cross-talk between the Stokes parameters which varies during the year between zero and 16%. Compensation of the telescopic birefringence using a Bowen compensator as done by Wiehr and Roßbach (1974) reduces this effect to a maximum of 2%.

The possibility to observe the Stoke profiles in real time allows

an easy adjustment of the Bowen compensator: The birefringence of the the uncompensated telescope partially converts the circularly into linearly polarized light. Hence, if one observes longitudinal magnetic fields with a *linear polarization modulator*, the characteristic signature of the V-Stokes profile is seen which disappears for a correctly adjusted Bowen compensator. The accuracy of this method was verified placing the telescope between crossed polarizers (Wiehr, 1971).

The strong influence of telescopic birefringence on the three Stokes profiles Q, U and V, is shown in Fig. 1 for both, uncompensated and compensated telescope. The occasionally observed double reversals in sunspot umbrae disappeared after fine-adjustment of the Bowen compensator. This suggests their origin in linear polarization via cross-talk.

Observations

In order to investigate the possibilities of the new apparatus, we observed several magnetic structures with various combinations of lines in May/June 1982 and 1983 at the Locarno observatory.

Among the numerous observations performed we consider as most interesting the asymmetry and the spatial extension of the Stokes-V profiles from solar network boundaries and from small plages. The selection of magnetic structures is done by means of the three-channel slit-jaw imaging device at the Locarno telescope (Wiehr et al. 1980).

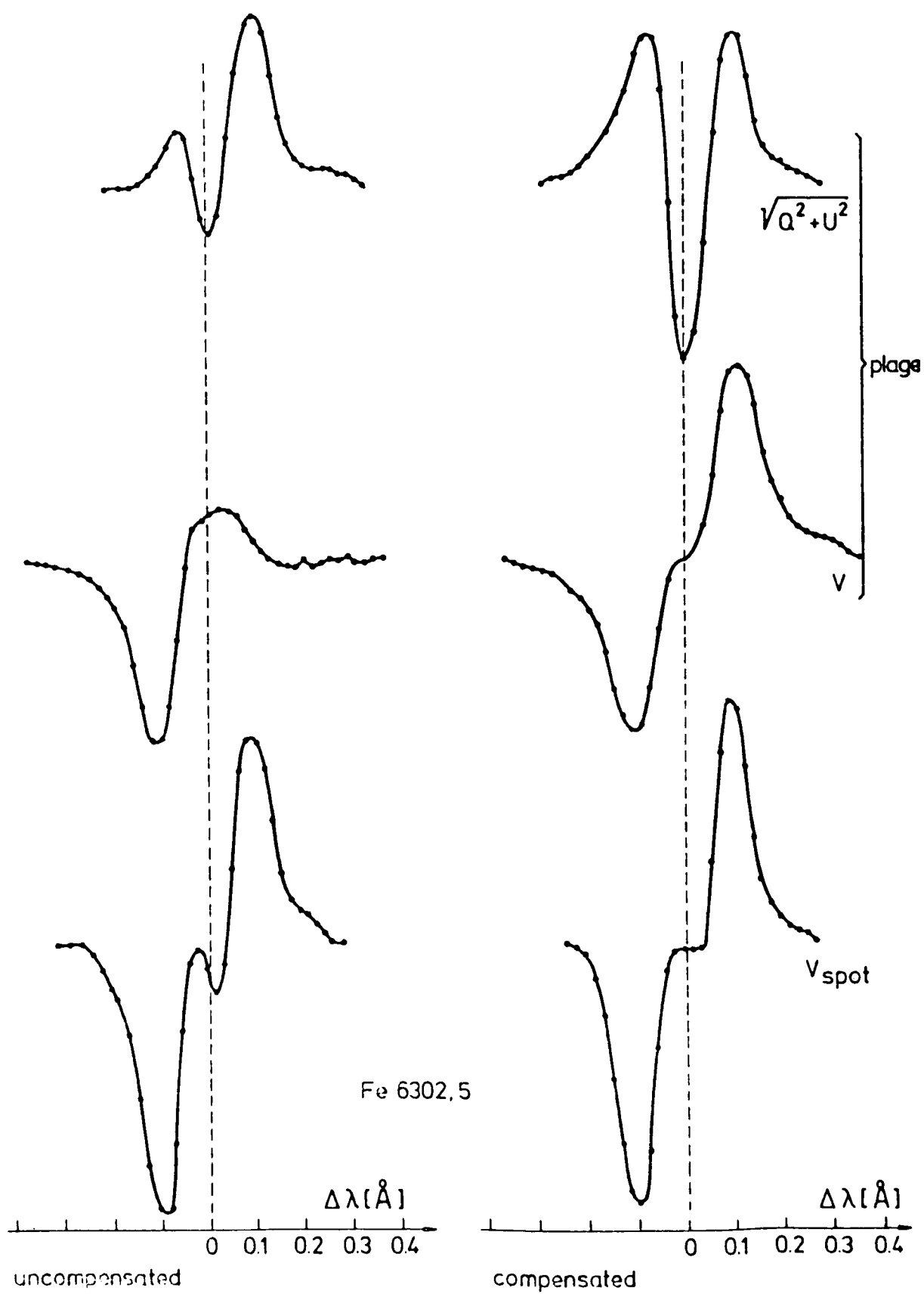


Figure 1

First results

Our extended measurements in network and plages at good spatial resolution show an asymmetry and an excess polarization of the blue side of the V-profiles. In order to quantify the asymmetry and the excess blue polarization, EBP, we determined the ratio of the peak polarization, $V_{\max}^{\text{blue}}/V_{\max}^{\text{red}}$, and that of the blue and red part of the total V-profile $\int V^{\text{blue}}/\int V^{\text{red}} \equiv \text{EBP}$, in network and plages.

Our results show that the peak ratios are almost the same, whereas the EBP values are found to be larger for the smaller split lines. The spatial variation of EBP in contrast to the almost constant peak ratios (Fig.2) indicates that at least two parameters are required for the description of the asymmetry of the V-profiles, and that spatial resolution is necessary for its detailed investigation. The variation of both parameters with the amplitude of the circular polarization, i.e. with the total magnetic flux, roughly confirms the findings by Stenflo and Harvey (1984) for the peak ratios obtained with the spectrometer, but not those from the F.T.S. data.

In contrast to the network and plage regions, our umbra V-profiles are highly symmetrical. We occasionally find a double reversal which disappears after compensation of the telescopic birefringence. In our case of V-double reversal a significant linear polarization always exists, and the central V-profile is asymmetric with respect to the zero line as well as to the total V-profile. If we try to interpret the hitherto observed double reversals in terms of a cross-talk between the Stokes parameters, we find support from Deubner and Liedler (1969) who observed it "particularly in the

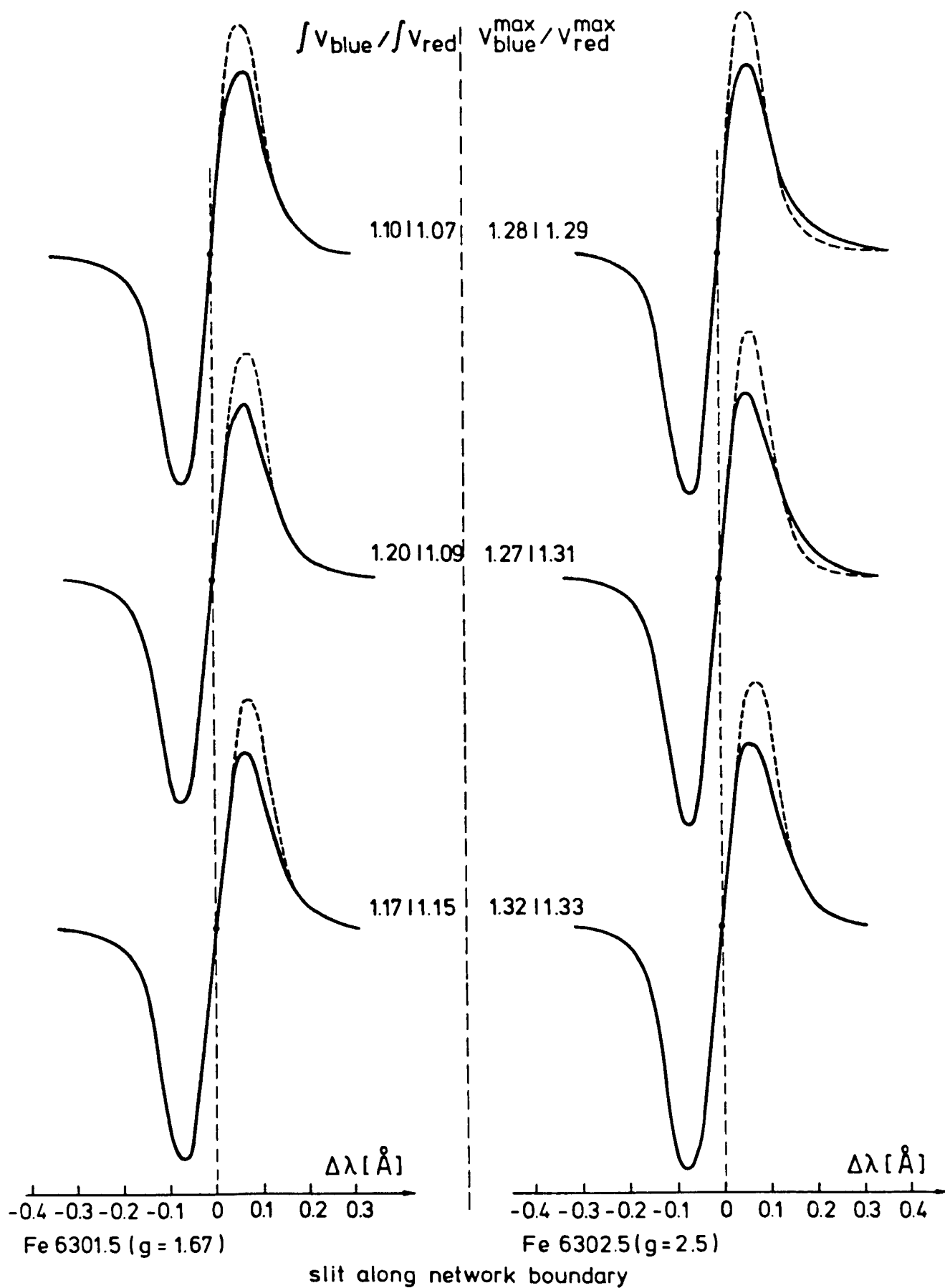


Figure 2

penumbra " where linear polarization certainly exists.

Simultaneous measurements of the V-Stokes parameter in the high-atmosphere Mgb 5167.3 and the neighbouring deep-atmosphere $\text{Fe}^+5169.1$ lines yield no measurable difference of the spatial extension of network magnetic fields covered by a 75 arcsec long slit. We estimate that the field lines do not open between the corresponding 0 km and, resp., 500 km levels by more than 1000 km.

Considering our measurements of all Stokes parameters in spots and plages, a severe problem arises from their fine structures which can not be resolved because of the rather long integration times required. In addition, the determination of the magnetic vector from the measured Stokes parameters encounters problems from the uncertain model atmospheres (e.g. Wiehr, 1974, §6), which are otherwise fairly well known for umbrae (e.g. Stellmacher and Wiehr, 1981a), for plages (e.g. Stellmacher and Wiehr, 1979) and for average penumbrae (Kjeldseth-Moe and Maltby, 1969). However, almost nothing is known about the actual influence of the fine structures on these models which, at least for the penumbrae where most transverse fields are measured, may be of essential importance (Stellmacher and Wiehr, 1981b).

References

- Deubner, F.L., Liedler, R.: 1969, Solar Physics 7, 87
Kjeldseth-Moe, K.O., Maltby, P.: 1969, Solar Physics 8, 275
Stellmacher, G., Wiehr, E.: 1979, Astron. Astrophys. 75, 263
Stellmacher, G., Wiehr, E.: 1981a, Astron. Astrophys. 95, 229
Stellmacher, G., Wiehr, E.: 1981b, Astron. Astrophys. 103, 211
Stenflo, J.O., Harvey, J.W.: 1984, Solar Physics (in press)
Wiehr, E.: 1971, Solar Physics 18, 226
Wiehr, E.: 1974, Solar Physics 95, 351
Wiehr, E., Roßbach, M.: 1974, Solar Physics 35, 343
Wiehr, E., Wittman, A., Wöhl, H.: 1980, Solar Physics 68, 207

ANALYSIS OF THE NEW POLARIMETER FOR THE
MARSHALL SPACE FLIGHT CENTER VECTOR MAGNETOGRAPH

E. A. West

Space Science Laboratory
NASA Marshall Space Flight Center
Huntsville, AL 35812

ABSTRACT

The Marshall Space Flight Center (MSFC) magnetograph has undergone an extensive upgrading in both electronic control of the magnetograph hardware and in the polarization optics. In the original MSFC magnetograph, two variable waveplates (KD*Ps) were used to measure the Stokes parameters. The problems associated with the original polarimeter were: 1. Field of view errors associated with the natural birefringence of the KD*P crystals, 2. KD*P electrode failure due to the halfwave DC voltage required in one of the operational sequences and, 3. Breakdown of the retardation properties of some KD*Ps when exposed to a zero to halfwave modulation (DC) scheme. To minimize these problems, one of the KD*Ps of the original polarimeter was replaced with a rotating waveplate assembly. Thus, the new polarimeter gives up the flexibility provided by two variable waveplates to adjust the retardances of the optics for a particular polarization measurement, but solves the problems associated with the original polarimeter. With the addition of the quartz quarterwave plates, a new optical alignment was developed to allow the remaining KD*P to correct for errors in the waveplates. In this paper, we present the new optical alignment of the polarimeter, discuss the various sources of error, and describe how those errors are minimized so that the MSFC magnetograph can look at the transverse field in "real time."

1. INTRODUCTION

One of the primary research goals of the Marshall Space Flight Center (MSFC) vector magnetograph is to study the evolution of vector magnetic fields over a large field of view. Figure 1 shows the optics box of the MSFC magnetograph with the various optical elements required to accomplish this goal. The optics box is mounted to a 12-inch cassegrain telescope which focuses the image of the sun onto a square aperture that limits our field of view of the Sun to 6X6 arc minutes. The light from the aperture is then collimated through the polarizing optics and Zeiss birefringent filter. After the Zeiss filter, a lens refocuses the light through the electronic shutter onto the CCD camera system.

The MSFC magnetograph has undergone two upgrades since 1981; one in the electronics and the second in the polarimeter. The electronics upgrade began with the replacement of the old SEC vidicon detector with a new solid state (CCD) camera system and ended with software control of hardware functions. The electronic upgrade will not be discussed in this paper other than to indicate that the signal to noise was improved from the original magnetograph's 100 to 1 per picture to ~ 800 to 1 per picture in the new magnetograph. Therefore some errors in the polarimetry, that were buried in the noise of the original magnetograph, would now require software corrections if improvements were not made in the optics.



FIGURE 1. OPTICS BOX OF THE MARSHALL SPACE FLIGHT
CENTER VECTOR MAGNETOGRAPH.

2. POLARIMETER ERRORS

Figure 2 compares the optical elements of the two polarimeters. In the original polarimeter, two KD*P (potassium dideuterium phosphate) crystals were used as variable waveplates. Since KD*P's are electro-optic crystals that allow rapid changes in retardation, they are important in minimizing seeing effects. In the original polarimeter, the fast axis of the first KD*P retarder was parallel to the Zeiss filter (which acted as the analyzer to the polarimeter); the fast axis of the second KD*P was at 45 degrees to the analyzer. In the new polarimeter, the first KD*P is replaced with a rotating waveplate assembly for the following reasons.

A. Temperature dependence

The temperature dependence of a KD*P crystal is approximately 1.75 degrees retardance per degree centigrade. To minimize the temperature variation in the optics box, the thermal control unit for the polarimeter is set to 5°C above the maximum expected temperature. If the temperature rises above the control unit setting, retardation errors occur. Removing one of the KD*Ps simplifies these errors so that software corrections can remove the crosstalk between the linear and circular polarization measurements. In the new magnetograph this problem is minimized further by the continuous monitoring of the temperature by the computer. When the temperature is out of a range, an alarm is set so that the observer can correct the problem.

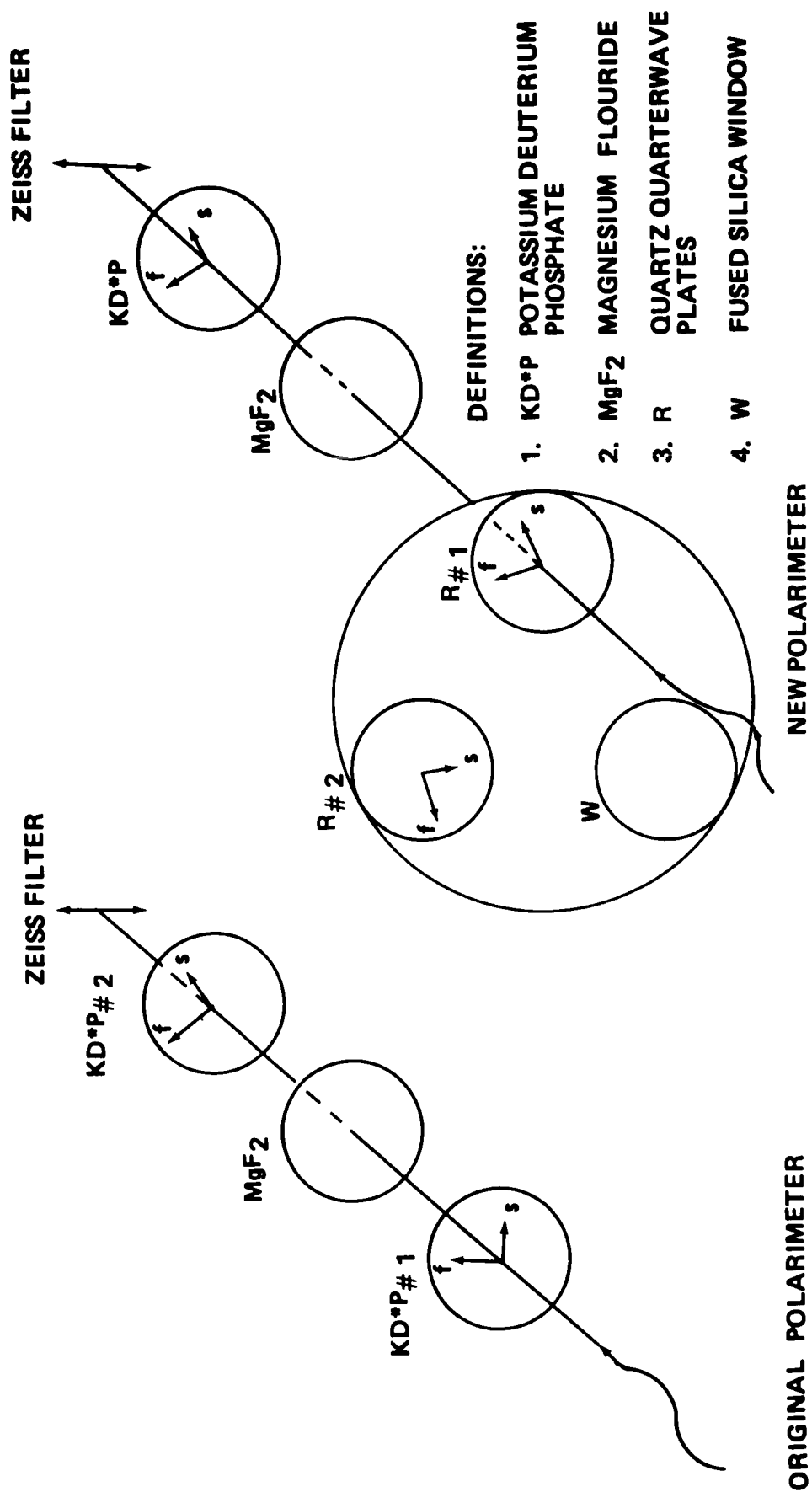


FIGURE 2. DIAGRAM DEPICTING THE OPTICAL ELEMENTS IN THE ORIGINAL AND NEW POLARIMETER.

B. Field of view errors

The second reason for replacing the first KD*P is related to the limited field-of-view of KD*Ps (West, 1978). KD*Ps are biaxial crystals when voltage is applied to them. This means that the retardation varies with the propagation direction of the light, with the retarding voltage and with the thickness of the KD*P. For example, the 6 arc-minute off axis error in retardation for one KD*P (whose thickness is approximately 2.5 mm) with a half wave voltage applied is approximately 8 degrees. In the original magnetograph, the second KD*P doubles the thickness which doubles the off axis errors. In the original polarimeter, a magnesium fluoride correction plate was added to the polarimeter to reduce by one half the field-of-view errors.

With only one KD*P in the new polarimeter, the field-of-view errors are equivalent to the best that the original polarimeter was capable of (~ 8 degrees). When a magnesium fluoride plate is added, the field-of-view errors are reduced by half (4 degrees). By replacing the first KD*P with two quarterwave plates and a window (the rotating waveplate assembly in Figure 2), the voltage applied to the remaining KD*P can be reduced by one half which has an equivalent effect on the field-of-view errors (2 degrees). Therefore the new polarimeter represents a significant improvement in the field-of-view errors over the original polarizing optics.

C. Crystal and electrode failures

The third problem area to be addressed by the new polarimeter is related to electrode and crystal failures that occurred with the unidirectional voltage modulation (0 to $1/2$ wave retardation) sequence in the original design. The replacement of the first KD*P with a waveplate wheel allows the retardation required to measure the Stokes vector with the remaining KD*P to be limited to plus and minus quarter-wave modulation, as indicated in Table 1. In the table S_A represents the signal detected by the CCD camera system when a plus $\lambda/4$ retardation (R_A) is applied to the remaining KD*P, S_B when a negative $\lambda/4$ retardation (R_B) is applied. The position of the waveplate then determines which element of the Stokes vector is measured. With the symmetric, "low" voltage (2500 volts) modulation, the KD*P lifetime has been increased substantially. For example, during 1980 while supporting the Solar Maximum Mission, the original polarimeter went through 5 KD*Ps. The new polarimeter is still using the same KD*P that was installed in June 1983.

In summary, the new polarimeter has simplified the thermal problems, has minimized retardation errors over the field of view and has extended the lifetime of KD*Ps. To accomplish these goals, the new polarimeter replaced a variable waveplate (KD*P) with waveplates that have fixed retardances. The accuracy of the new polarimeter is dependent upon the quality of these waveplates.

WAVE PLATE POSITION

**S_A = SIGNAL WHEN $+\lambda/4$
ON KD*P (R_A)**

**S_B = SIGNAL WHEN $-\lambda/4$
ON KD*P (R_B)**

$\frac{S_A - S_B}{S_A + S_B}$ = % POLARIZATION

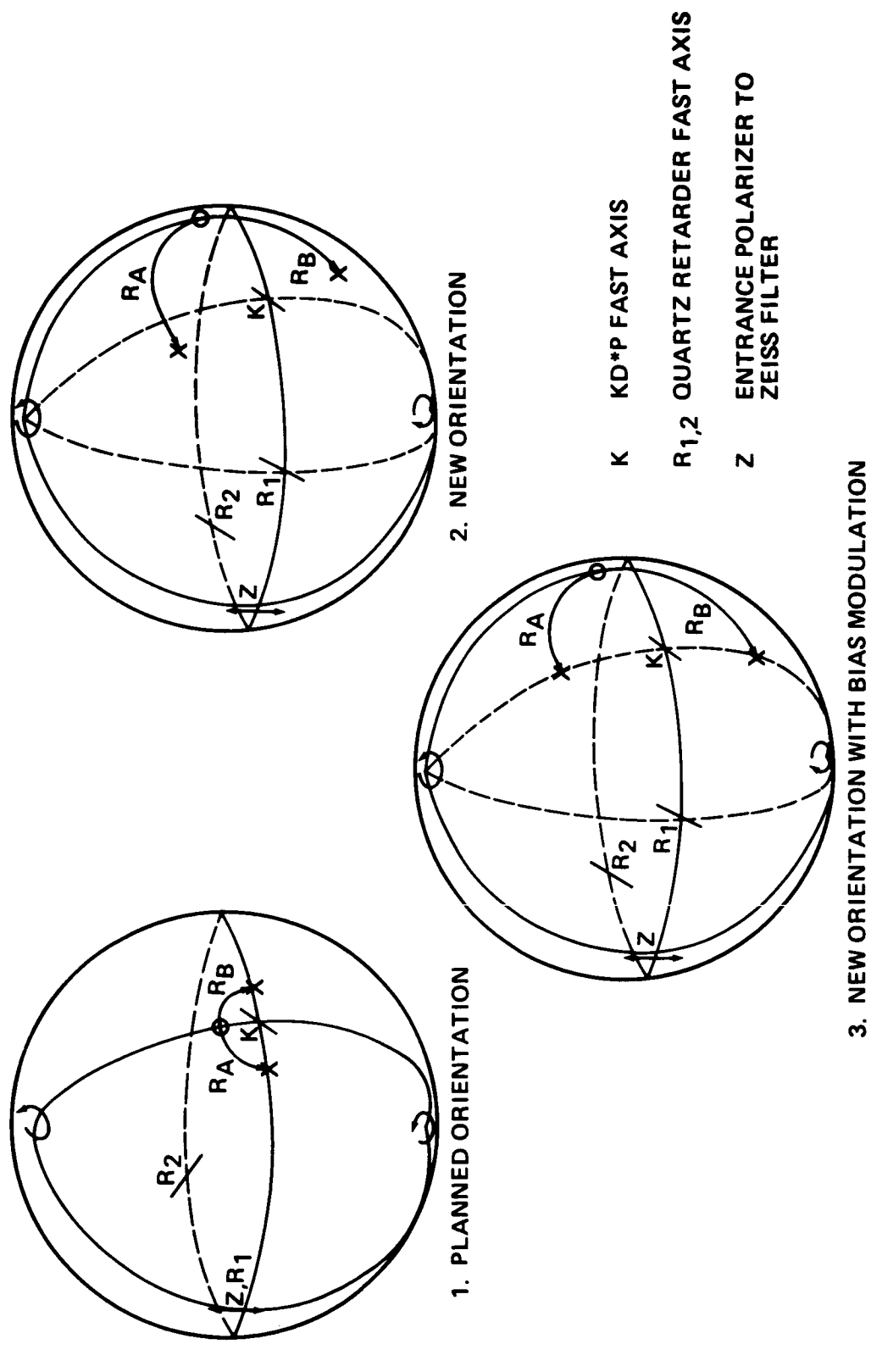
I. WINDOW	II. RETARDER # 1	III. RETARDER # 2
I + V	I + Q	I + U
I - V	I - Q	I - U
V/I	Q/I	U/I

**TABLE 1. RELATIONSHIP BETWEEN WAVEPLATE POSITION AND THE MEASURED STOKES VECTOR
(I, Q, U, V).**

3. QUARTERWAVE PLATE ERRORS

The natural evolution in the design of the new polarimeter was to assume that the fast axes of the quarterwave plates would have the same orientation as the optical elements in the original polarimeter: the fast axis of one retarder aligned parallel to the fast axis of the remaining KD*P and the other parallel to the entrance polarizer of the Zeiss filter (Figure 2). Although the quarterwave plates had tight tolerances on flatness, parallelism and retardance, the most difficult tolerance to obtain by the manufacturer (and to verify at MSFC) was the retardance. The specification for the waveplate retardance at 525.0 nm was 90 ± 1 degree. The measured retardance for the two waveplates were 91.2 and 88.8 with an accuracy of ± 2.5 degrees. If either waveplate had been a "perfect" 90 degree retarder, the planned orientation would have been successful. Unfortunately, difficulties in minimizing circular crosstalk in both waveplates when aligned parallel to the Zeiss filter required a new orientation.

Figure 3 compares the Poincare sphere representation of three modulation schemes for the MSFC polarimeter. The measured linear polarizations (Q and U) are defined to be 45 degrees from the retarder fast axis positions (R_1 and R_2). In the planned orientation, the wave plate R_1 would rotate circular polarization to linear polarization parallel to the fast axis of the KD*P; then any voltage (retardation) applied to the KD*P would produce no change in the signal level transmitted through the Zeiss filter (Z on Poincare sphere). If the waveplate has a retardation error and stops short of the linear axis



O LOCATION OF CIRCULAR POLARIZATION AFTER 85 DEGREE ROTATION BY RETARDER R1

X FINAL LOCATION OF CIRCULAR POLARIZATION AFTER KD*P ROTATION IN RA OR RB DIRECTION

FIGURE 3. POINCARÉ SPHERE REPRESENTATION OF THREE MODULATION SCHEMES FOR THE MSFC POLARIMETER

(indicated by small "o" on Poincare sphere), then circular crosstalk will get into the linear measurements as demonstrated by the difference in the final positions of R_A and R_B (Figure 3, planned orientation). To allow the remaining KD*P to minimize circular crosstalk a new orientation had to be considered. In the new orientation the waveplates are placed at ± 22.5 degrees from the Zeiss filter. If plus and minus quarterwave voltages are applied to the KD*P (90 degrees retardance is approximately 2200 volts DC), the new orientation will have the same circular crosstalk as the planned orientation (1 and 2 in Table 2; Figure 3, new orientation). But if the KD*P voltages (retardation) have a DC bias, the circular crosstalk in the new orientation can be eliminated (Figure 3, new orientation with bias modulation). This is demonstrated in Table 2. Although the bias modulation eliminates circular crosstalk it will produce linear crosstalk (3 in Table 2). In solar measurements, linear crosstalk is not as critical as circular (V), but can be eliminated by fine tuning the fast axes of the two waveplates until the measured linear polarizations are 45 degrees apart.

	RETARDER (R ₁) ORIENTATION	KD*P RETARDATION		STOKES PARAMETERS %		
		RA	RB	Q/I	U/I	V/I
1.	PLANNED	90	-90	99.6	0	8.7
2.	NEW	90	-90	99.6	0.0	8.7
3.	NEW (BIAS)	83	-97	99.6	8.7	0.0

Q - LINEAR POLARIZATION RELATIVE TO R₁

U - LINEAR POLARIZATION RELATIVE TO R₂

V - CIRCULAR POLARIZATION

I - INTENSITY

TABLE 2. COMPARISON OF THE MEASURED STOKES PARAMETERS FOR THE THREE MODULATION SCHEMES OF THE MSFC POLARIMETER ASSUMING A 5 DEGREE ERROR IN THE RETARDER R₁.

4. SUMMARY

In conclusion, the performance of the polarimeter has been improved by simplifying the thermal problems, minimizing the field-of-view errors and reducing the KD*P failures. With waveplates in the correct orientation, the circular crosstalk that is induced with imperfect retarders can be minimized with the remaining KD*P. Finally, the linear crosstalk induced by the odd KD*P modulation can itself be reduced by fine tuning the positions of the fast axes of the quarterwave plates.

REFERENCES

West E. A., 1978, Applied Optics, 17, p 3010.

THE SOLAR VECTOR MAGNETOGRAPH OF THE OKAYAMA ASTROPHYSICAL OBSERVATORY

Mitsugu Makita, Shigeo Hamana and Keizo Nishi

Tokyo Astronomical Observatory, University of Tokyo

Abstract: The vector magnetograph of the Okayama Astrophysical Observatory is fed to the 65 cm solar coude telescope with a 10 m Littrow spectrograph. The polarimeter put at the telescope focus (solar image size = 31 cm) analyzes the incident polarization. Photomultipliers (PMT) at the exit of the spectrograph pick up the modulated light signals and send them to the electronic controller. The controller makes frequency and phase analyses of the signal, and the results are digitized and stored on the magnetic disk of the mini-computer.

The analyzer of the polarimeter is a combination of a single wave plate rotating at 40 Hz and a Wallaston prism. Incident linear and circular polarizations are modified at four times and twice the rotation frequency, respectively. Two compensators minimize the instrumental polarization, mainly caused by the two tilt mirrors in the optical path of the telescope. The observation is performed stepwise from point to point on the sun with a photoelectric guider in the side telescope, controlled by the mini-computer. The over-all accuracy is better than 0.1 percent polarization for an integration time of 1 second observing in the green spectral region.

The four photomultipliers placed on the wings of the FeI 5250A line give maps of intensity, longitudinal field (circular polarization), and transverse field (linear polarization). A line shifter in front of the PMTs affords us a map of the velocity field. Besides these main outputs, maps of intensity, and net linear and circular polarizations in the neighboring continuum (5282A-5297A) are obtained by the other two monitor PMTs. An action region (8 arc min x 8 arc min) is typically swept in about 70 minutes, and the maps are obtained in 15 minutes after finishing the observation.

1. Introduction

The plan to build a vector magnetograph in the Okayama Astrophysical Observatory started by measuring the instrumental polarization (Makita and Nishi, 1970) and proposing its compensation (Makita, 1970). The first financial support was given in 1975 and the polarimeter was manufactured by Nikon. Since then, necessary instrumental units were supplied one by one almost every year. The total system was thus completed by obtaining the mini-computer and color graphic tube in 1982. The operation of the complete vector magnetograph became possible in autumn 1982. Some preliminary data mainly on solar active regions were published recently (Sakurai and Makita, 1984).

2. Construction of the Vector Magnetograph

The vector magnetograph of the Okayama Astrophysical Observatory is installed on the 65 cm solar coude telescope. Figure 1 shows the construction of the instruments schematically. The solar image as large as 31 cm in diameter is formed at the coude focus. The polarimeter is put there and analyzes the

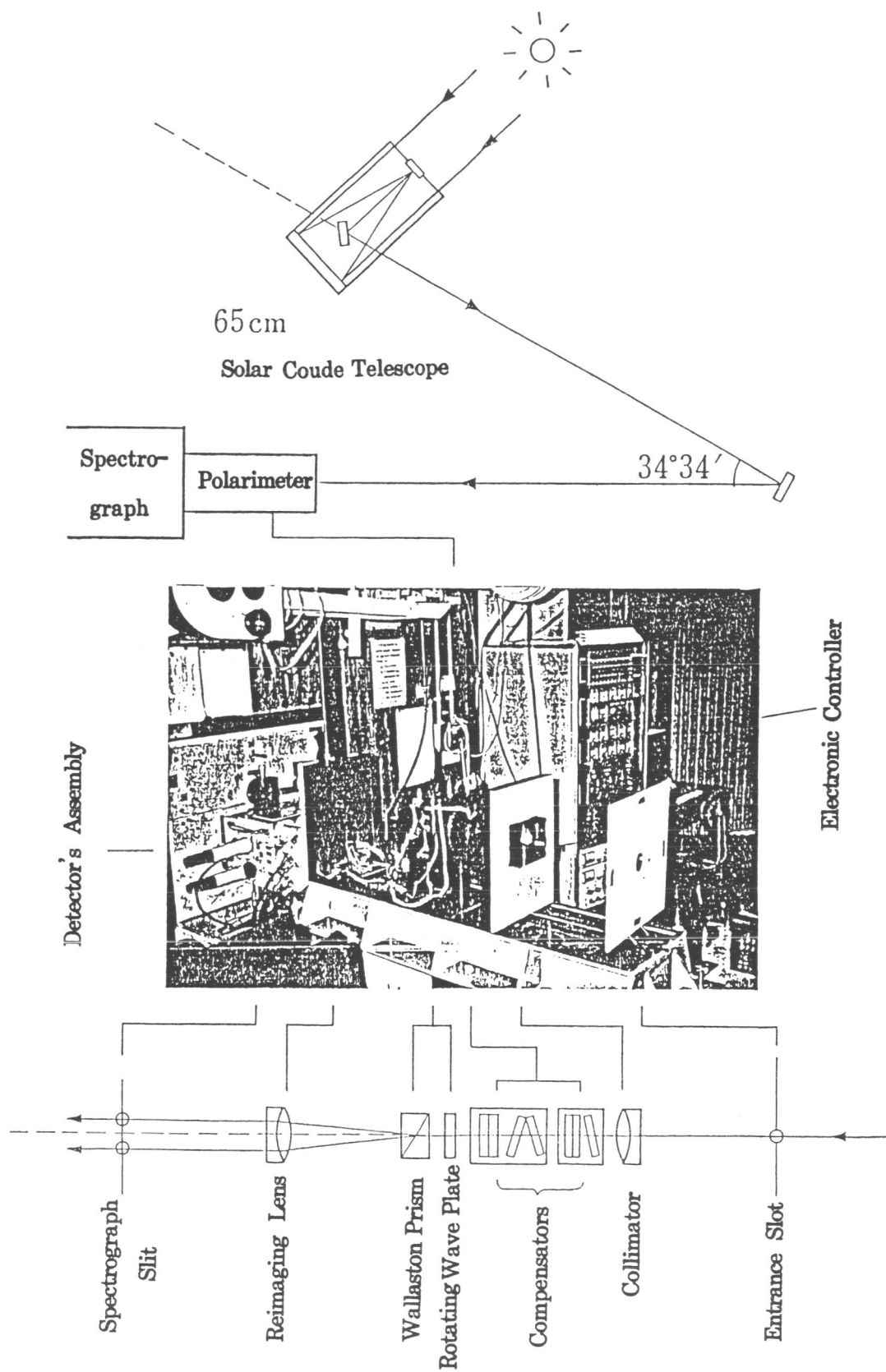


Fig. 1. Solar vector magnetograph of the Okayama Astrophysical Observatory.

polarization of the incident light beam passing through its entrance slot. The light beam in the polarimeter is collimated, passes through compensators for the instrumental polarization, then is separated by the analyzer and reimaged on the spectrograph slit. The spectrograph is a Littrow type with 10 m focal length. At the exit of the spectrograph six photomultipliers (PMT; EMI6256B) detect the output light signals. Figure 2 shows this schematically. Two pairs of PMTs pick up the signals from the wings of a Zeeman sensitive line, and the other pair monitors the neighboring continuum. A line shifter in front of the spectrograph balances the intensity signals from the wings A and B, and gives the radial velocity. Accordingly, six polarization signals and one velocity signal are sent to the magnetograph electronic controller. This makes frequency and phase analyses of the polarization signals and amplification. The resultant outputs are digitized in 12-bits and stored on the magnetic disk of the mini-computer, MELCOM 70/60. Some characteristics of the instrument are described in the following section.

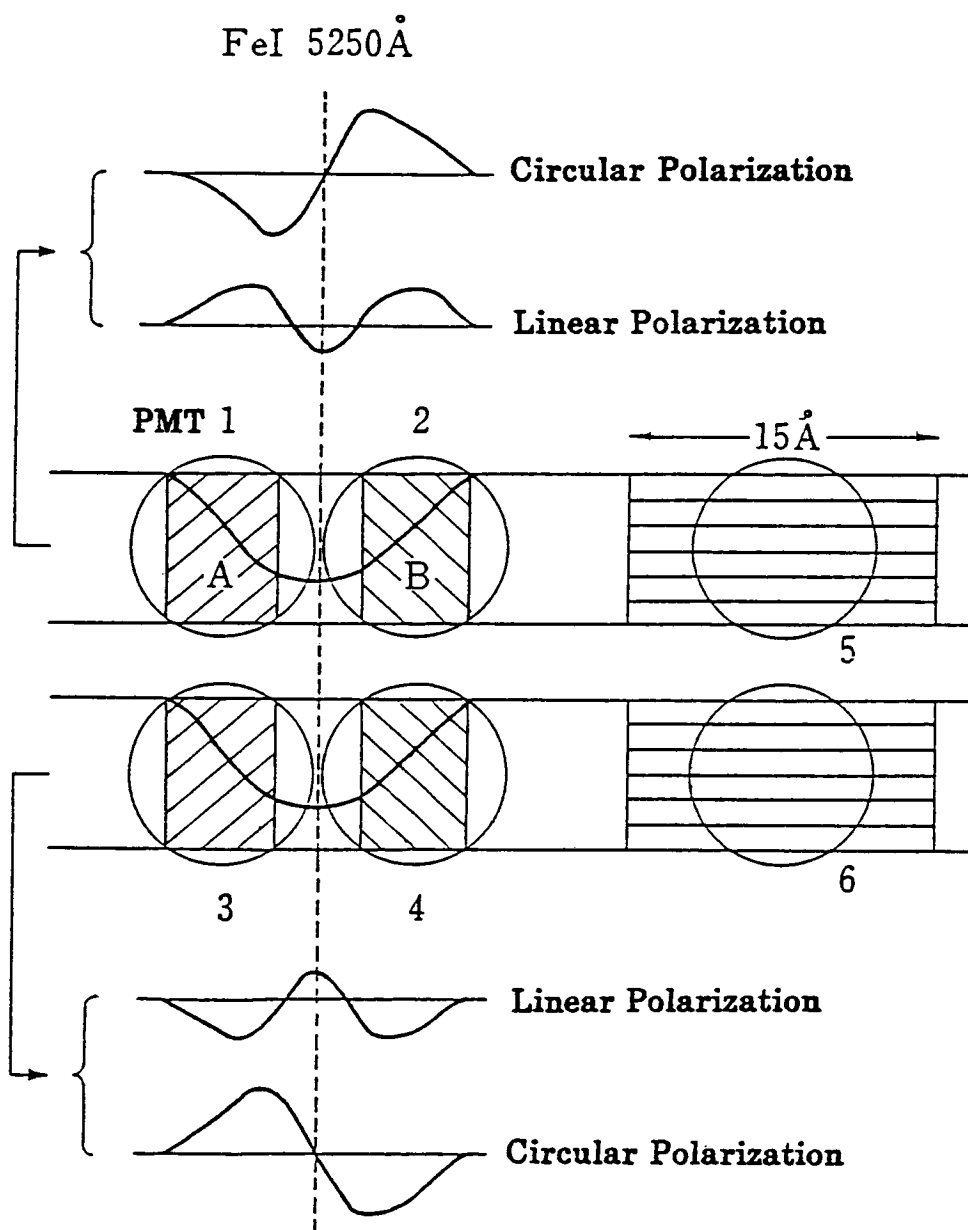


Fig. 2. Arrangement of the photomultipliers (PMT) and polarization signals of the spectral line.

3. Some Characteristics of the Vector Magnetograph

A. Rotating Wave Plate

The polarimeter has been investigated in detail in previous work (Makita et al., 1982). Therefore, the analyzer, the main component of the polarimeter, is briefly described here. It is the combination of a wave plate rotating at 40 Hz and a Wallaston prism. This modulates incident polarization as follows (Makita et al., 1982).

$$I_{\text{obs}} = \frac{1}{2} [I + \frac{1}{2} (1 + \cos\delta) Q + \frac{1}{2} (1 - \cos\delta) (Q \cos 4\Omega t + U \sin 4\Omega t) + V \sin\delta \sin 2\Omega t + R \sin(\delta_{\perp} + \delta_{\parallel}) \sin\delta \cos 2\Omega t] ,$$

where (I, Q, U, V) are the Stokes parameters of the incident light, δ is the phase retardation due to the wave plate, Ω the angular frequency of the rotating wave plate, R the amplitude of the component caused by the surface reflection, and δ_{\perp} and δ_{\parallel} are the phase change of the linearly polarized light perpendicular and parallel to the optic axis of the wave plate. Since the Stokes parameters are modified with different frequencies and phases, they can be completely separated and obtained by frequency and phase analysis for an arbitrary but known retardation, δ . The rotating wave plate, as a polarization analyzer, was proposed by Sekera (Hodgdon, 1965), and actually used, for example, by Akabane (1958) to observe solar radio emission and by Orrall (1971).

Interference fringes due to reflections at surfaces of the wave plate are clearly visible in the frequency and phase analysis as shown in Figure 3 (see also the formula above). Temperature drift of these fringes may affect the measurement of the radial velocity. A thin wave plate as thick as 100 microns broadens the fringe separation as wide as 9λ near 5250Å, far wider than the spectral line width, and makes the drift rather insensitive to the velocity measurement.

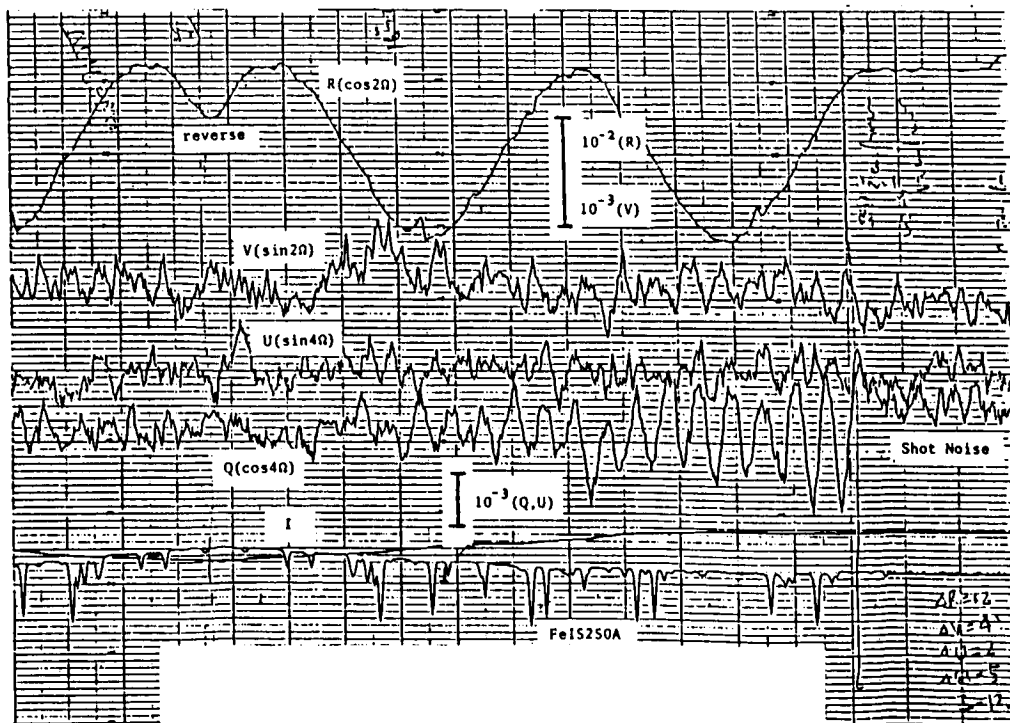


Fig. 3. Spectram scanning of the polarization signals (I, Q, U, V, R) and shot noise in the quiet region.

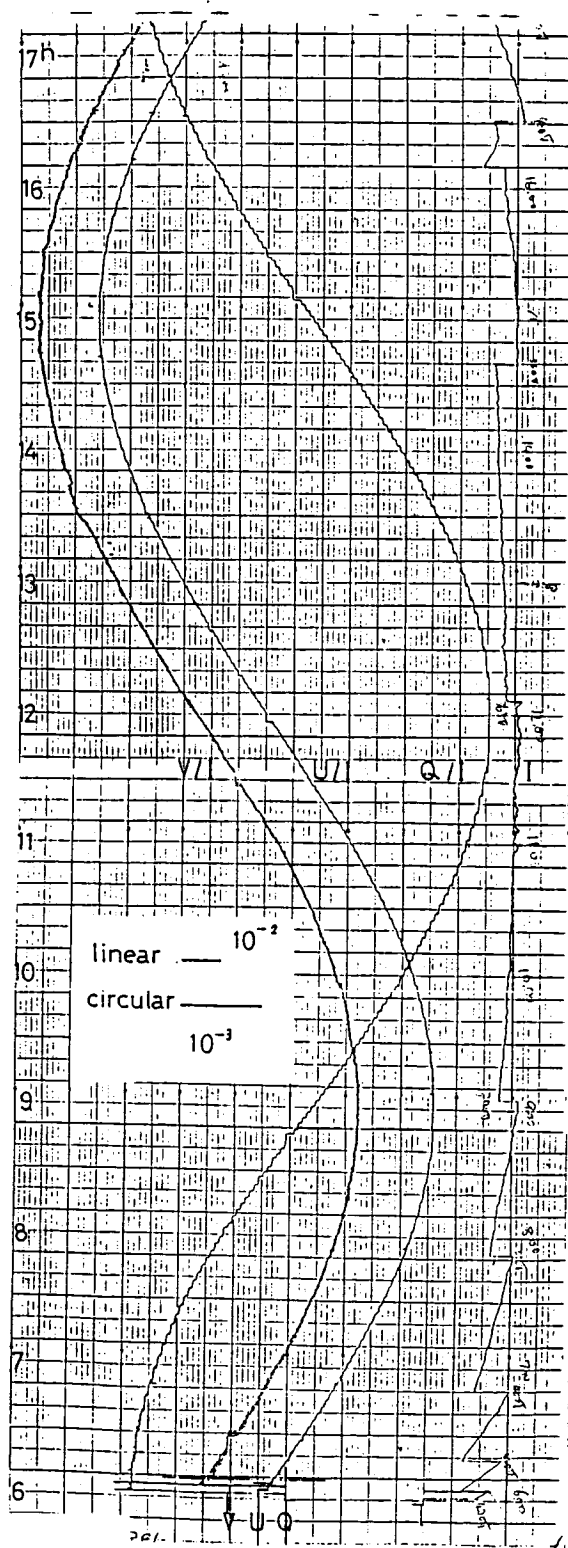


Fig. 4. Diurnal variation of the instrumental polarization of the solar coude telescope.

Electronic crosstalk caused by the reflection component, R , is minimized by an anti-reflection coating on the wave plate. A slight wedge shape of the wave plate initially made for non-uniform transmission and interference with the internally reflected rays which, coupled with the non-axisymmetric light beam of the telescope, produced considerable erroneous signals. This was removed fairly well by polishing the wave plate surfaces to be as parallel as possible.

B. Compensators

The instrumental polarization measured by the polarimeter is reproduced in Figure 4 (Makita et al., 1982). The disk center of the sun was observed from morning to evening. The unpolarized light is linearly polarized by the first tilt mirror, and a small part of this linear polarization is transformed into circular polarization by the second tilt mirror (See Figure 1). The diurnal variation is due mainly to the rotation of the first mirror in the telescope tube, since the reflection angles at the two mirrors remain almost constant during the observation. The degree of polarization seasonally changes as the reflection angle of the first mirror changes. Its amplitude reaches 6×10^{-2} or 2×10^{-3} in summer, for the linear or circular polarization. These are far greater than the expected observing accuracy of $10^{-4} \sim 10^{-3}$.

A compensator to eliminate the large instrumental polarization was proposed by Makita (1970). Two compensators are put in front of the analyzer. The first one is for the second mirror, and the next one, diurnally rotating around the optical axis, is for the first mirror. Each compensator is composed of plane parallels and a Babinet compensator. The tilt of the plane parallel compensates the linear polarization caused by the reflection. Two plane parallels instead of a single one are installed in the rotating compensator to make the tilts smaller and the adjustment by the tilt variation easier (Makita 1970). The Babinet compensators are composed of two wave plates, the optic axes of which are orthogonal. The tilt of one of the components compensates the instrumental phase retardation, changing the path length of the light beam. This is not the ordinary Babinet compensator, but a simplified one.

A serious problem was surface reflections of each component, which produce linearly polarized spectral fringes (e.g. Makita et al., 1982). To overcome this, we tried to make each component a lambda plate, thus producing no surface reflection for the observing wavelength. Figure 3 shows for Q-polarization the spectral fringes almost certainly due to the non-rotating compensator, which has 17 lambda wave plates as thick as 0.9692 mm. The rotating compensator, having 59 lambda wave plates as thick as 2.9656 mm, does not give remarkable fringes. All the components in the compensators are coated to minimize surface reflections in the observing wavelengths. The automatic compensation proposed in the earlier paper (Makita, 1970) was not performed since good compensation continued during the observation period.

C. Calibration and Accuracy

Calibration of the polarization degree was made by tilt of a plane parallel with and without a quarter wave plate for the linear and circular polarizations, respectively. The amount of the tilt is well calibrated, and the polarization degree corresponding to this amount is calculated and compared with the output signal (Makita et al., 1982). Dividing factors of the linear and circular polarization signals are 6.1 and 12.4 at 5250A, and 3.4 and 6.1 in the neighboring continuum, 5282A-5297A.

The overall accuracy for spectral line observations is seen in Figure 3. The shot noise and, in the case of Q-polarization, the spectral fringes near the observing line have at most amplitudes of 0.1 per cent, when the integration time is one second. The signals from the monitor PMTs show the far smaller fluctuation as 0.01 per cent (Makita et al., 1982).

The transformation of the degree of polarization to the magnetic field strength has not been done. The limited data, however, show that the circular polarization of 4×10^{-3} corresponds to about 5G in Mount Wilson magnetograms and the linear polarization of 7.5×10^{-3} to about 200G in the Sib-IZMIR magnetograms.

3. Data Acquisition and Display

Observations are usually made by using the Zeeman sensitive line FeI 5250A and the neighboring continuum 5282A-5297A, simultaneously. The retardation of the rotating wave plate is $7/4$ lambda and equivalent to minus $1/4$ lambda at 5250A. The same rotating wave plate can be used in other wavelength regions by reinterpreting the retardation (see the formula in section 3A).

The observing area is defined by the slot of the polarimeter and the slit of the spectrograph. In the typical case this is 12 arc sec x 0.6 arc sec, the latter being equivalent to the slit width of 100 microns. Observing points are selected stepwise by a photoelectric guider in the side telescope under the control of the minicomputer. Typical scanning steps are 10 arc sec in east-west and north-south directions. Observations take at least 70 minutes to measure 50 x 45 points, when the integration time is one second and the scan step is 10 arc sec. A video camera monitors in H-alpha the slot diaphragm picture of the polarimeter and can record it during the scanning.

It is a difficult problem to select in which part of the line profile the exit slits of the spectrograph should be placed. There have been some reports about this (Severny 1964; West and Hagyard, 1983; Kawakami, 1983). The present exit slits cover the spectral band and 27 mA to 80 mA relative to line center. With these exit slits, directions of the linear polarization closely agreed with those of the net linear polarization of a sunspot obtained by the monitor PMTs. The inner part of the profile such as 5 mA-48 mA gave the discrepancy as shown in Figure 5. This will be discussed elsewhere.



Fig. 5. Directions of the linear polarization in an active region observed on May 25, 1983. Dotted lines are obtained with the exit of 0.027-0.080Å at UT 0h25m-1h23m and solid lines with the exit of 0.005-0.048Å at UT 1h31m-2h26m. A closed curve shows a sunspot.

DATE 84: 04: 13 TIME 01: 00: 38 TO 02: 03: 10
 LINEAR POLARIZATION

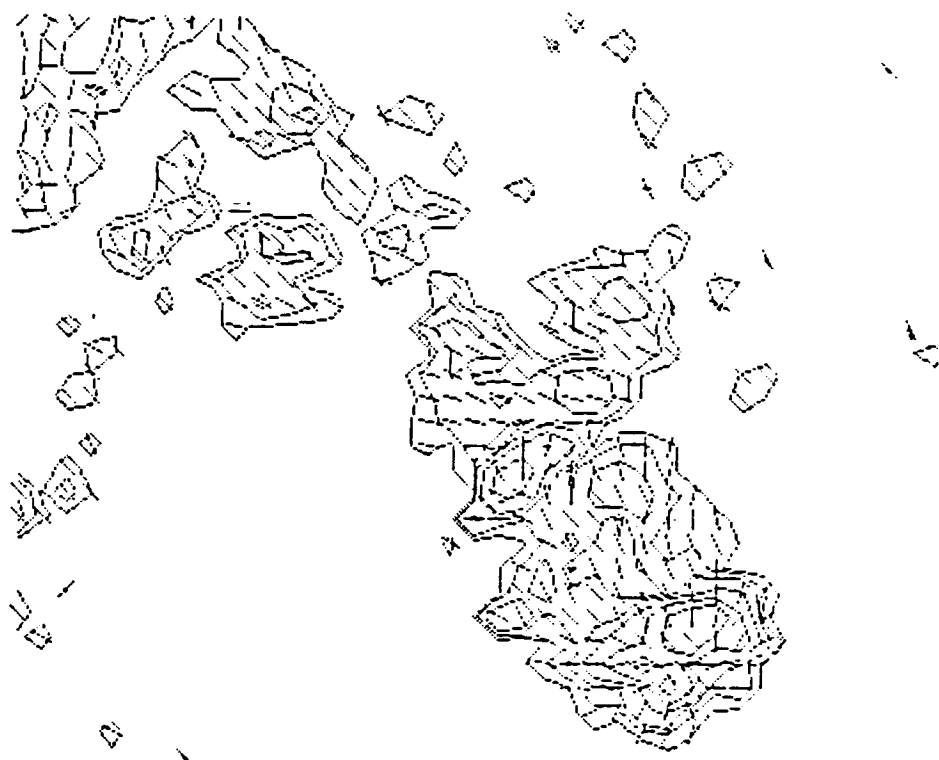


Fig. 6a. Linear polarization of an active region. Polarization degree is 2.46×10^{-3} for the outermost curves and increases to the inner ones in a geometrical progression with the equal ratio of 2. Bars show the direction of the polarization. Dark square on the lower right corner is 10×10 arcsec and roughly equal to the spatial resolution.

DATE 84: 04: 13 TIME 01: 00: 38 TO 02: 03: 10

CIRCULAR POLARIZATION



Fig. 6b. Circular polarization of an active region. Polarization degree is 4.03×10^{-3} for the outermost curves and increases to the inner ones in a geometrical progression with the equal ratio of 2. Solid and dashed curves show the north and south magnetic polarity, respectively.

DATE 84: 04: 13 TIME 01: 00: 38 TO 02: 03: 10

VELOCITY FIELD

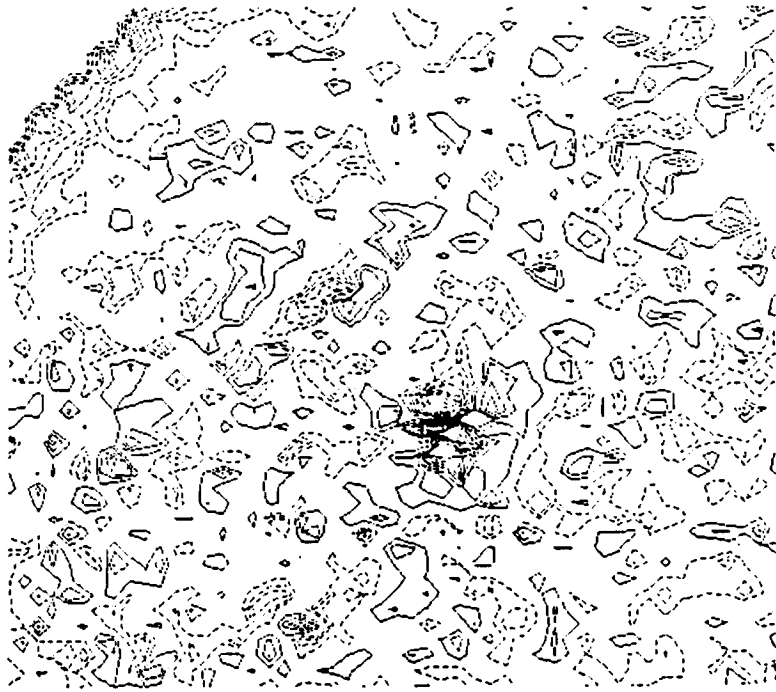


Fig. 6c. Radial velocity in an active region. Radial velocity is 110 m/s for the outermost curves and increases stepwise to the inner ones by the same amount. Solid and dashed curves show the blue and red shifts, respectively.

Results of the observation temporarily stored on the magnetic disk are the Stokes parameters and Doppler shift of the spectral line and the Stokes parameters of the neighboring continuum. During the next observation they are transferred to magnetic tape for the detailed analysis. A quick display of the observed data can be made on the color graphic tube by using the temporary record. Figure 6 gives an example of the line polarization and velocity maps obtained with the attached copier. All the above data processing is completed within 15 minutes. In Figure 7, color graphic displays of analyzed data are shown.

Acknowledgements

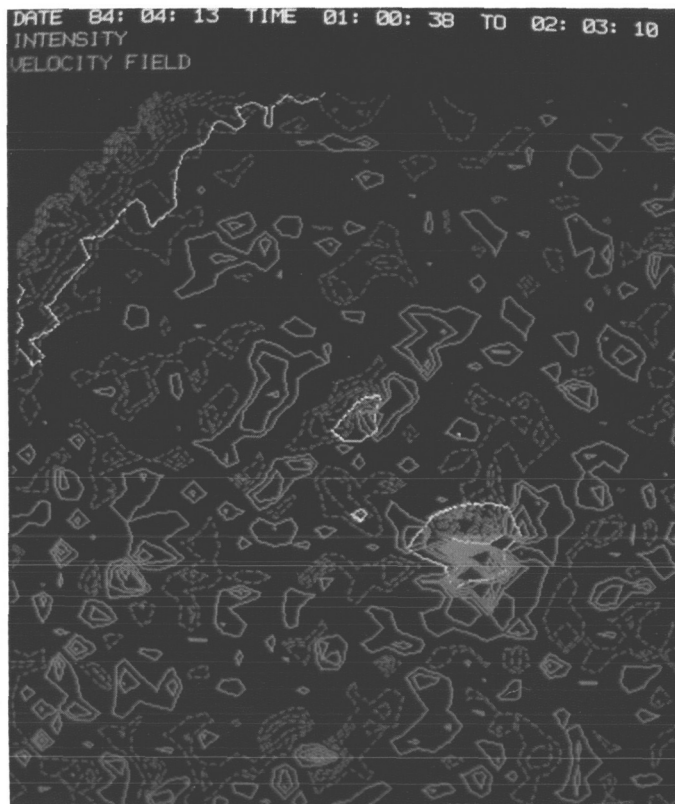
Thanks are due to Mr. M. Shimizu and the staff of the Okayama Astrophysical Observatory for the assistance during the experiments and observations. Thanks are also due to Dr. T. Sakurai of Department of Astronomy, University of Tokyo, and Dr. K. Shibasaki of Research Institute of Atmospheric Sciences, Nagoya University, for improving the data display program. The polarimeter and the magnetograph electronic controller were made by Nikon. The authors very much appreciate the cooperation of the Oyo Kodensho Co. in developing the wave plates for the analyser and compensators.

References

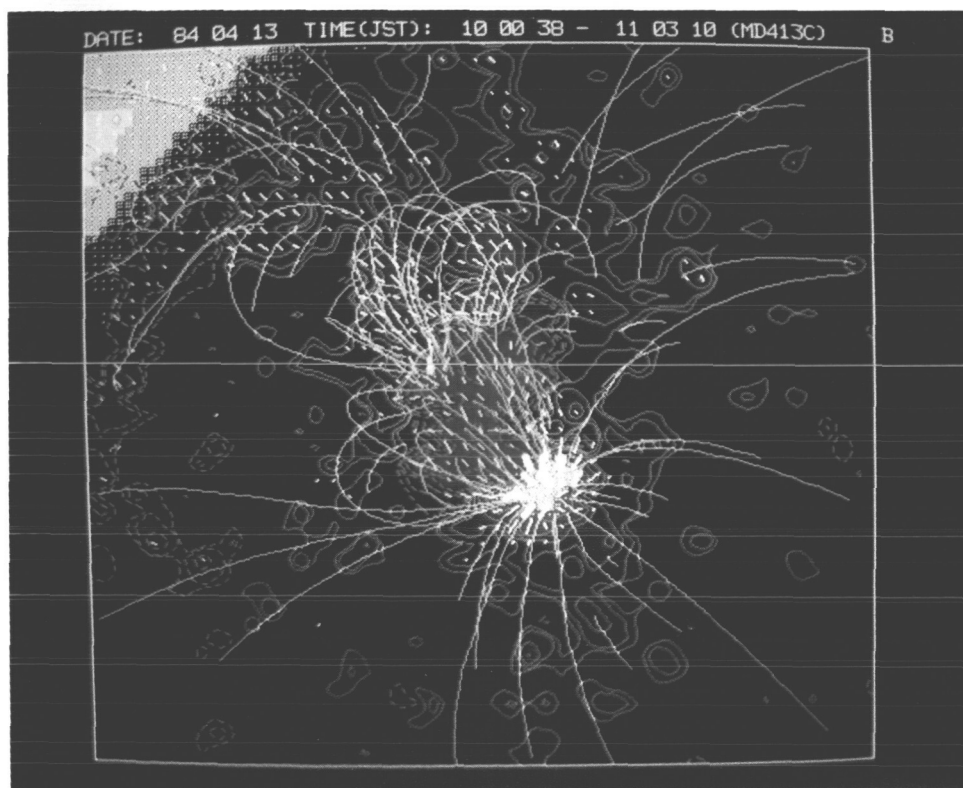
- Akabane, K.: 1958, *Annals Tokyo Astron. Obs.*, 6, 57.
Hodgdon, E. B.: 1965, *Applied Optics*, 4, 1479.
Kawakami, H.: 1983, *Publ. Astron. Soc. Japan*, 35, 459.
Makita, M. and Nishi, K.: 1970, *Annals Tokyo Astron. Obs.*, 12, 121.
Makita, M.: 1970, *Annals Tokyo Astron. Obs.*, 12, 139.
Makita, M., Hamana, S., Kawakami, H. and Nishi, K.: 1982, *Annals Tokyo Astron. Obs.*, 19, 24.
Orrall, F. Q.: 1971, *Solar Magnetic Fields*, ed. R. Howard, IAU Symp. No. 43, p. 30.
Sakurai, T. and Makita, M.: 1984, *Vector Magnetograms of Solar Active Regions*, December 1982-December 1983, Tokyo Astronomical Observatory.
Severny, A. B.: 1964, *Izv. Krim. Astrofiz. Obs.*, 31, 126.
West, E. A. and Hagyard, M. J.: 1983, *Solar Phys.*, 88, 51.



(A)



(B)



(C)

Figure 7. Observations of the magnetic and velocity fields of a solar active region in April 1984. (a) Superposition of linear polarization (green line segments) representing the transverse magnetic field on the circular polarization (colored contours) representing positive (blue) and negative (red) line-of-sight fields. (b) The velocity field of the active region. Blue and red contours denote positive and negative line-of-sight velocities; green contours show the intensity distribution in the sunspots. (c) Potential field calculation. Field lines derived from a potential theory calculation are shown superposed on the line-of-sight magnetic field.

VECTOR MAGNETIC FIELD OBSERVATIONS WITH THE HALEAKALA POLARIMETER

D. L. Mickey

University of Hawaii, Institute for Astronomy
Haleakala Observatory, P. O. Box 209, Kula, Hawaii 96790

Abstract. Several enhancements have recently been made to the Haleakala polarimeter. Linear array detectors provide simultaneous resolution over a 3-Å wavelength range, with spectral resolution of 40 mÅ. Optical fibers are now used to carry the intensity-modulated light from the rotating quarter-wave plate polarimeter to the echelle spectrometer, permitting its removal from the spar to a more stable environment. These changes, together with improved quarter-wave plates, have reduced systematic errors to a few parts in 10^4 for routine observations. Examples of Stokes profiles and derived magnetic field maps are presented.

1. INTRODUCTION

The University of Hawaii has operated a solar Stokes polarimeter since around 1972, in a gradually evolving configuration. Originally, the instrument was configured for low-spectral-resolution, high-sensitivity observations using interference filters and photomultiplier detectors. Later, an echelle spectrometer and silicon diode detectors were added, permitting line profile observations over a wider spectral range but with reduced sensitivity. Recent enhancements have resulted in significant reductions in systematic errors and in increased throughput, thereby greatly increasing the range of observational programs possible with this instrument. I would like here to present an overview of the Haleakala polarimeter, highlighting the recent improvements and indicating its special capabilities (and limitations). I'll show a "vector magnetogram" derived from Stokes profile observations, then present a few of the profiles themselves to indicate the character of the basic data and suggest some areas for caution in interpretation. Finally, I will suggest a few types of observations, for which this instrument is especially well qualified, that have promise of adding to our capabilities for understanding solar vector magnetic field observations.

2. INSTRUMENT DESCRIPTION

A schematic of the Haleakala polarimeter is shown in Figure 1; its main features are listed in Table 1. The on-axis coronagraph telescope minimizes both scattered light and instrumental polarization. A pinhole in the focal plane selects the field of view; due to diffraction losses, pinholes smaller than 50-μm diameter (angular field 4 arc sec) are not very useful. Behind the pinhole the beam is collimated and passes through a rotating wave-plate modulator and Wollaston prism analyzer. The wave plate currently in use is a two-element quartz device, optically contacted to minimize the number of surfaces and keep beam deviation below one or two arc seconds. The intensity of light emerging from the analyzer is given by

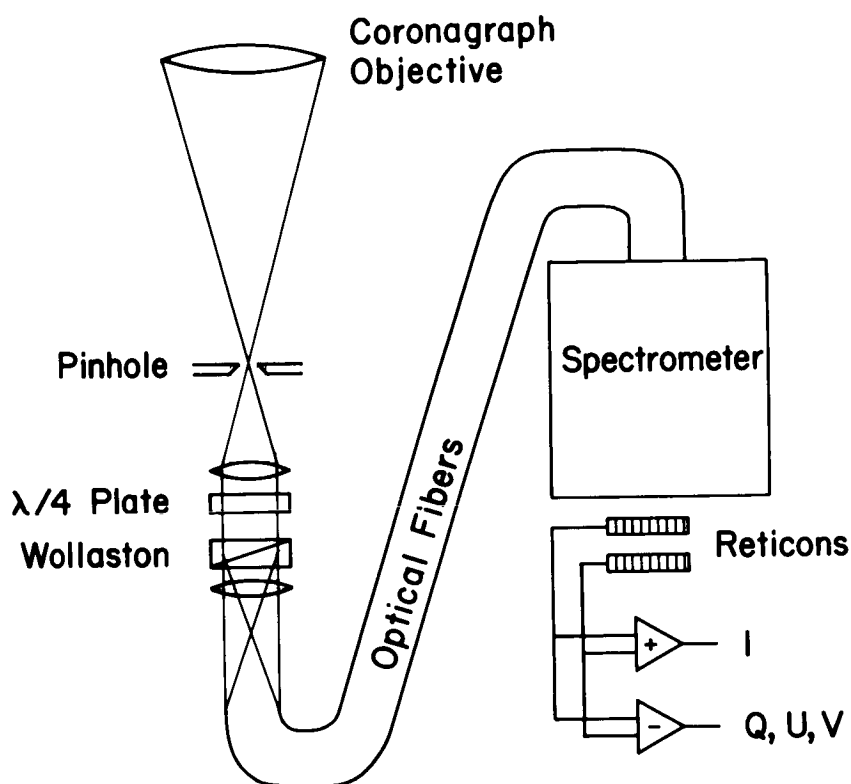


Fig. 1. Optical schematic of the Haleakala Stokes polarimeter.

$$I' = \frac{1}{2} \left[I \frac{Q}{2} \pm (1 + \cos \delta) \mp V \sin \delta \sin 2\theta \right. \\ \left. \pm \frac{Q}{2} (1 - \cos \delta) \cos 4\theta \pm \frac{U}{2} (1 - \cos \delta) \sin 4\theta \right], \quad (1)$$

where I , Q , U , V are the Stokes parameters of the incoming beam, δ is the retardance of the wave plate, and θ is its rotation azimuth, counterclockwise direction from the $+Q$ direction. The signs on the Q , U , V terms refer to the two beams emerging from the analyzer. The precise value of δ is not critical in the range $0.25-0.4 \lambda$; the wave plate currently in use has a retardance of 207 nm. A problem we have encountered with extremely flat wave plates involves interference effects due to multiple reflections within the plate: for light polarized either parallel or perpendicular to its fast axis, weak interference fringes appear with a spacing of typically 1-2 Å. Because the plate has different optical thicknesses for the two polarizations, the two sets of fringes are displaced from each other by roughly half their spacing. This causes the wave plate to exhibit a fringe pattern in its linear polarizance, $P(\lambda)$, which in turn adds a spurious $P(\lambda)I \cos 2\theta$ term to Equation 1.

TABLE 1

PRINCIPAL COMPONENTS OF THE HALEAKALA STOKES POLARIMETER

Telescope

15-cm aperture, f:16
 Coronagraph lens
 On-axis design
 Spatial resolution set by pinhole field stop (4, 6, 8, 10
 arc sec diameter)

Polarimeter

Rotating $\lambda/4$ plate modulator optically contacted quartz,
 $\lambda/4$ at 830 nm
 Wollaston prism analyzer
 Both beams utilized

Fiber-Optic Feed

80 μm x 15 m long quartz fibers

Spectrometer

1-m echelle, f:11
 Asymmetric Czerny-Turner configuration
 Two beams

Detectors

128-element Reticon linear arrays
 Scale 2.5 μm pixel⁻¹

The amplitude of this term can be on the order of 0.1 I for an uncoated plate, so it is necessary both to use high-performance AR coatings on the plate and to carefully adjust the zero of θ to prevent crosstalk between the $\cos 2\theta$ and $\sin 2\theta$ terms.

A major change made to the optical configuration during the past year was to insert a fiber-optic coupler between the polarimeter and the spectrometer. Previously, the spectrometer was mounted directly to the telescope, with two images of the pinhole formed directly on the slit. In that arrangement, however, even small amounts of wedge in the wave plate caused the images to wobble on the entrance slit, introducing spurious--and not very stable--intensity modulation at various frequencies. I have now removed the spectrometer from the telescope and located it in a stable environment in an adjacent room. Light from the two analyzer beams is carried to the spectrometer by a pair of quartz fibers that are 80 μm in diameter by approximately 15 m long. The thorough scrambling of the light within the fibers provides two important benefits: the two beams entering the spectrometer are now unpolarized, and--since the image at the entrance to the fiber is smaller than the fiber diameter--intensity modulation due to beam wobble is eliminated. The fibers in use (Fujikura S80-125) introduce very little beam spreading: light enters the fiber at f:7, emerges at f:6 or so, then is imaged onto the spectrometer slit at f:10 to match the spectrometer optics. The overall transfer efficiency of the fiber

system is about half that of the previous arrangement. I should point out that since the spectrometer slit now samples the image of the fiber exit face and the pinhole image is completely scrambled within the fiber, the spatial field is determined only by the pinhole, rather than the projection of the slit on the pinhole as before.

The spectrometer used in this system is a 0.75-m focal length Czerny-Turner design with a 316 l/mm echelle grating. Coma correction is achieved by an asymmetrical configuration of the collimator and camera mirrors, but the remaining astigmatism and spherical aberration of the full system limit its resolving power to 1.6×10^5 .

Another significant change made recently to the Haleakala polarimeter was the addition of linear array detectors. I have incorporated the 128-element Reticon arrays and associated electronics formerly used in the Stokes II system at Sacramento Peak (generously made available by the High Altitude Observatory). The detector package is essentially identical to that described by Baur *et al.* (1981); incorporation into the spectrometer at Haleakala required only re-imaging the spectrum to an appropriate scale and some software development. The image scale obtained was 2.5 μm per pixel at 630 nm, which provides adequate if not excessive sampling.

A summary of the instrument's capabilities is presented in Table 2. Although its spatial and temporal resolution are modest, its spectral resolution, freedom from scattered light and polarization sensitivity make it uniquely qualified for careful Stokes profile observations.

TABLE 2

MAJOR OBSERVATIONAL PARAMETERS OF THE HALEAKALA POLARIMETER

<u>Spatial Resolution:</u>	4-10 arc sec
<u>Spectral resolving power:</u>	160,000
<u>Time resolution:</u>	3 seconds per spatial point (disk, 8 arc sec field, λ 630 nm)
<u>Sensitivity:</u>	10^{-3} or better
<u>Scattered light:</u>	low, both in telescope and spectrometer
<u>Accessible wavelengths:</u>	530-880 nm (could be extended)

3. OBSERVATIONS

The derivation of solar magnetic fields from Stokes profile observations is a process for which neither the methodology nor the meaning of the result is as yet well understood. We want, nevertheless, to obtain "vector magnetograms" from our data, in order to study evolution of active regions, spot formation, and so on. We have used observed Stokes profiles to obtain maps such as that in Figure 2. This figure represents a raster scan of AR 4474, taken on 1984 April 30. Both the field stop and the raster step size were 8 arc sec; the integration time was 4 seconds per point. These maps are based on integrated Stokes profiles in the following way:

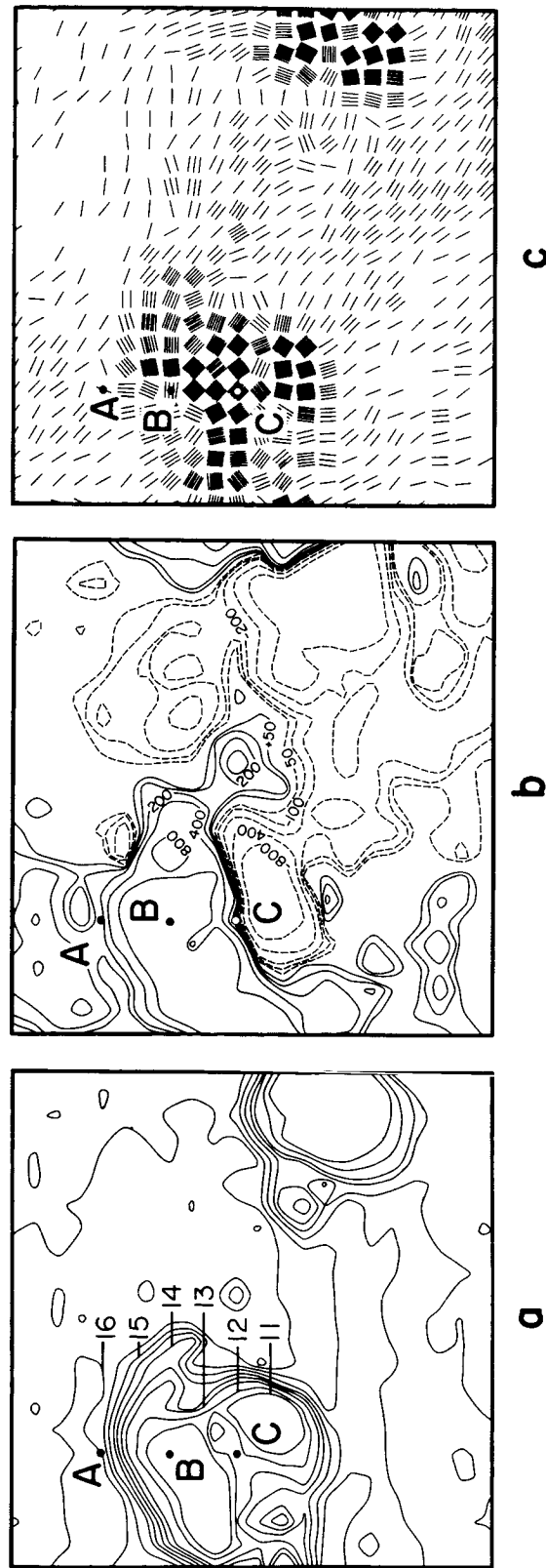


Fig. 2. Vector magnetic field maps of AR 4474 on 1984 April 30 at 2100 UT. (a) Brightness, arbitrary scale. (b) Line-of-sight component of B. (c) Transverse component of B (each line in a group represents 100 G). Stokes profiles for points marked A, B, and C are shown in Figs. 3 and 5.

The full width at half maximum, w , of the quiet-Sun intensity profile is determined, then integrals from $-1.5 w$ to $+1.5 w$ are taken for all four profiles at each raster point. We have

$$\langle I \rangle = \int_{-3w/2}^{+3w/2} I(\lambda) d\lambda, \quad (2)$$

$$\langle Q \rangle = \int_{-3w/2}^{-w/3} Q(\lambda) d\lambda - \int_{-w/3}^{+w/3} Q(\lambda) d\lambda + \int_{+w/3}^{+3w/2} Q(\lambda) d\lambda, \quad (3)$$

$$\langle V \rangle = \int_{-3w/2}^0 V(\lambda) d\lambda - \int_0^{+3w/2} V(\lambda) d\lambda, \quad (4)$$

and $\langle U \rangle$ is obtained in the same way as $\langle Q \rangle$. Then

$$B_L = K_1 \langle V \rangle / \langle I \rangle \quad (5)$$

and

$$B_T = K_2 ((\langle Q \rangle^2 + \langle U \rangle^2)^{1/2} / \langle I \rangle)^{1/2}, \quad (6)$$

where K_1 and K_2 are empirically determined calibration factors.

Repeating a raster like this shows rather good reproducibility for transverse fields above 100 G or so and longitudinal fields above a few gauss. It should be stressed, however, that these integrals are proportional to field strength only when the magnetic splitting is small compared to the line width, and that as pointed out by Stenflo (1984), the transverse component can be severely exaggerated if the magnetic elements comprise only a fraction of the field of view.

Figure 3 shows a typical set of Stokes profiles from the same data set, taken at the point marked A. All plots are in the same (arbitrary) intensity units. The derived field values here are $B_L = 200$ G, $B_T = 200$ G. The asymmetry and absence of a zero crossing in Q and U are not at all unusual. A point with nearly vertical (line-of-sight) field from the umbra at B is shown in Figure 4. For interest, and to encourage caution in the use of these magnetograms, I present Figure 5, from the point marked C in Figure 2. Although one cannot completely rule out instrumental contamination in the V profile, this pattern occurs consistently enough to suggest its origin is solar.

Another case of this type is shown in Figure 6, from a different sunspot. In both instances, the field is nearly transverse, and the V profile behaves as if there are two components separated by a velocity shift of around 8 km s^{-1} . Clearly, though the "magnetograph" approach given above would not be a complete failure for these lines, some of the most interesting physics is contained in the details of the line profiles.

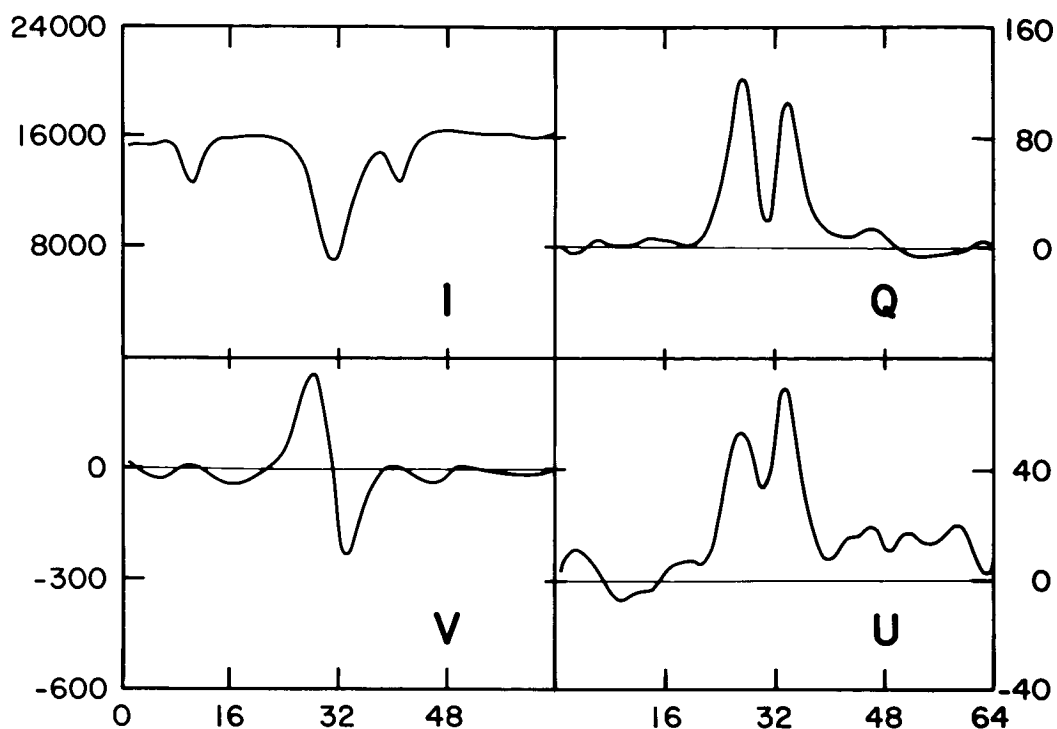


Fig. 3. Stokes profiles of Fe I $\lambda 6302.5$, from point marked A in Fig. 2. All profiles are in (arbitrary) intensity units. Q and U have not been rotated to a preferred frame.

4. OPPORTUNITIES

Several ideas present themselves for capitalizing on the capabilities of this instrument. For example, the low scattered light in the telescope should permit some prominence observations, including multi-line measurements as suggested by Bommier (1984). Further study of the situations where the lines give a non-zero net circular polarization in sunspots would be of interest, as would investigations of magneto-optic effects and "strange" profiles. If observational biases due to limited spatial resolution can be understood, we can perhaps take advantage of multi-line observations to build up a more accurate three-dimensional picture of magnetic fields in sunspots.

REFERENCES

- Baur, T. G., Elmore, D. E., Lee, R. H., Querfeld, C. W., and Rogers, S. R.: 1981, *Solar Phys.* **70**, 395.
 Bommier, V.: 1984, these proceedings.
 Stenflo, J. O.: 1984, these proceedings.

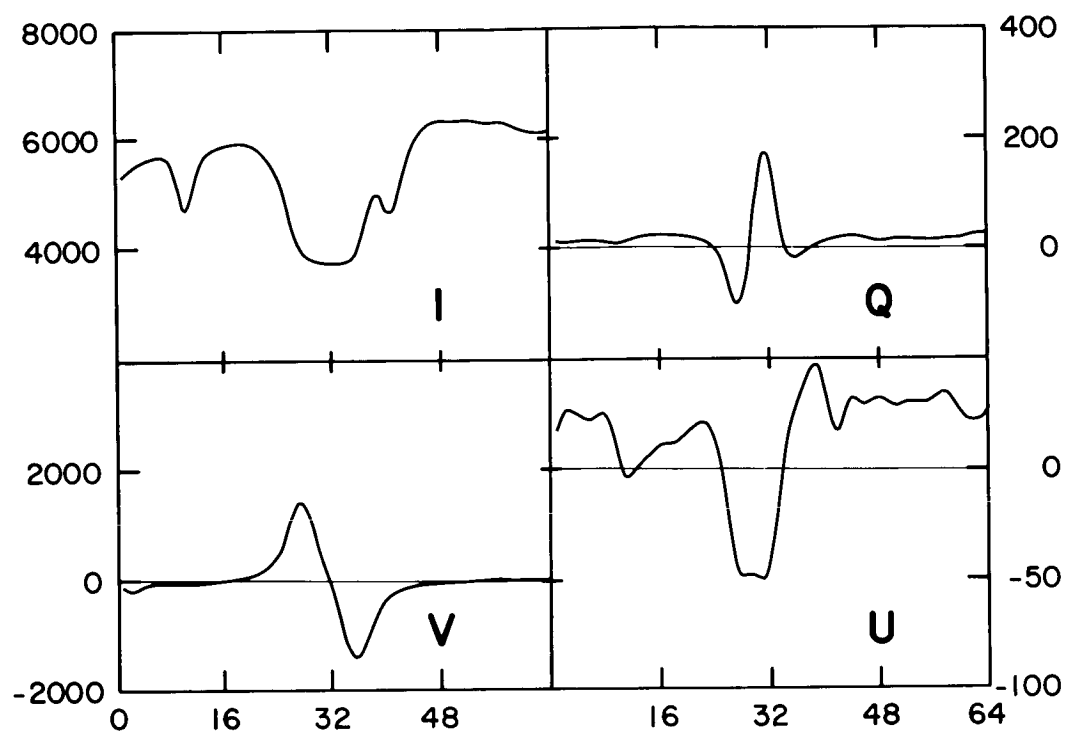


Fig. 4. $\lambda 6302.5$ profiles from point B in Fig. 2.

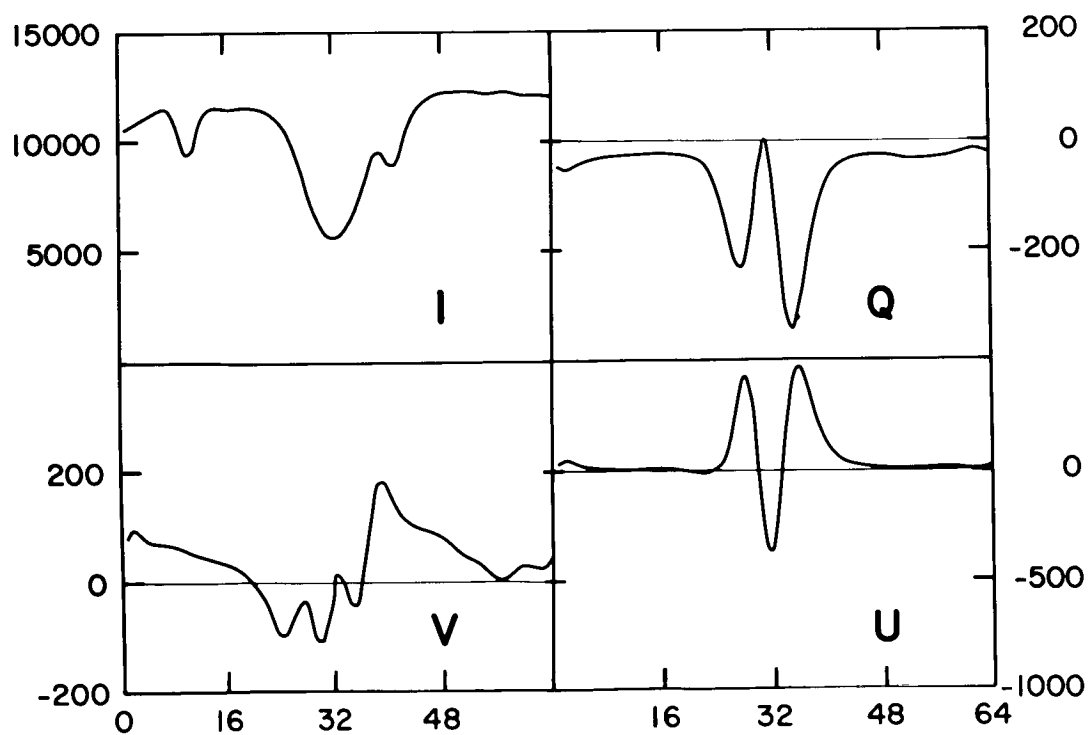


Fig. 5. $\lambda 6302.5$ profiles from point C in Fig. 2.

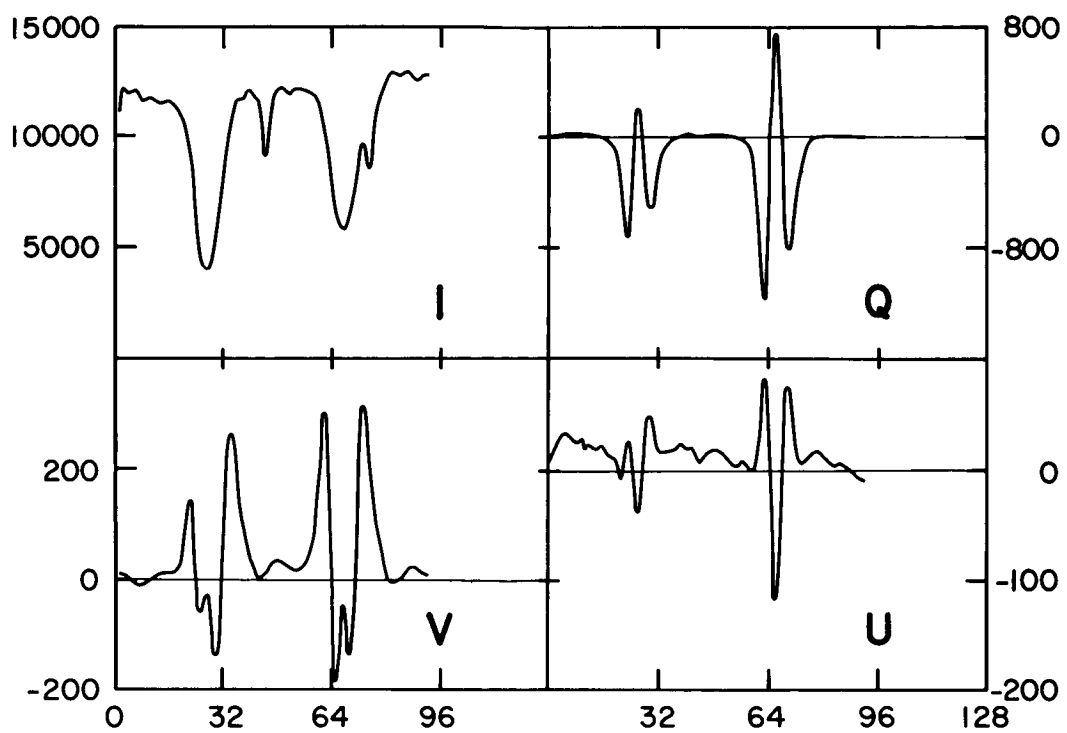


Fig. 6. $\lambda 6301.5$ and $\lambda 6302.5$ profiles from a spot penumbra on 1984 April 18.

STOKES PARAMETERS MODULATOR FOR BIREFRINGENT FILTERS

Audouin Dollfus

Observatoire de Paris, 92195 Meudon, France

Abstract. The Solar Birefringent Filter (Filter Polarisant Solaire Selectif FPSS) of Meudon Observatory is presently located at the focus of a solar refractor with a 28 cm lens directly pointed at the Sun. It produces a diffraction limited image without instrumental polarization and with a spectral resolution of 4.6×10^4 in a field of 6 arc min. diameter. The instrument is calibrated for absolute Doppler velocity measurements and is presently used for quantitative imagery of the radial velocity motions in the photosphere. We record the short period oscillations, the larger scale convection velocities and the motions in sunspots, usually with the lines F_eI 5577, NaD_1 , $H\alpha$ and $Mg\ b_1$. Our best radial velocity images have a resolving power of 500 km at the solar surface.

We are working to adapt the instrument for the imagery of the solar surface in the Stokes parameters. The first polarizer of the birefringent filter, with a reference position angle 0° , is associated with a fixed quarter wave plate at $+45^\circ$. A rotating quarter wave plate is set at 0° and can be turned by incremented steps of exactly $+45^\circ$. Another quarter wave plate also initially set at 0° is simultaneously incremented by -45° but only on each even step of the first plate. A complete cycle of increments produces images for each of the 6 parameters $I \pm Q$, $I \pm U$ and $I \pm V$. These images are then subtracted by pairs to produce a full image in the three Stokes parameters Q , U and V .

With proper retardation tolerance and positioning accuracy of the quarter wave plates, the cross-talk between the Stokes parameters was calculated and checked to be minimal. The two achromatic 36 mm diameter quarter wave plates are made of quartz and MgF_2 , and each held in a 8 position Genova-cross mounting. Proper gears associated with the drive shafts of the two Genova-crosses produce relative rotation rates of $+1$ and $-1/2$. Each of the six Stokes parameters can be selected at will, or they can be sequentially driven at a maximum rate of several positions each second.

1. The Birefringent Filter FPSS at Observatoire de Meudon

The Universal birefringent filter FPSS (for "Filtre Polarisant Solaire Selectif") was designed at Meudon Observatory for high angular resolution imaging of the solar surface with a FWHM selectivity of 130 mÅ at $\lambda = 5896$ (D_1 line). The linear field diameter is 36 mm. An entrance beam convergence of $f/30$ is accepted with an angular resolution that is diffraction limited for a convergence of $\leq f/45$. In order to increase the transparency, four of the eight polarizers used in the instrument are birefringent calcite-glass Rochon prisms of 96% transmission, the others being Polaroid HN 38 sheets of transmission 76%. The overall transmission produced by the polarizers together is $1/2 \times (0.96)^4 \times (0.76)^4 = 0.14$. The detailed description of the FPSS filter has not yet been published but technical reports are available at Meudon Observatory.

The FPSS instrument is presently used at Meudon at the $f/60$ focus of a 28 cm solar refractor directly pointed at the Sun (Fig. 1). The lens happens to be practically free of birefringence so that there is no interaction with the polarization properties of the incident beam, at the threshold of the detection level.

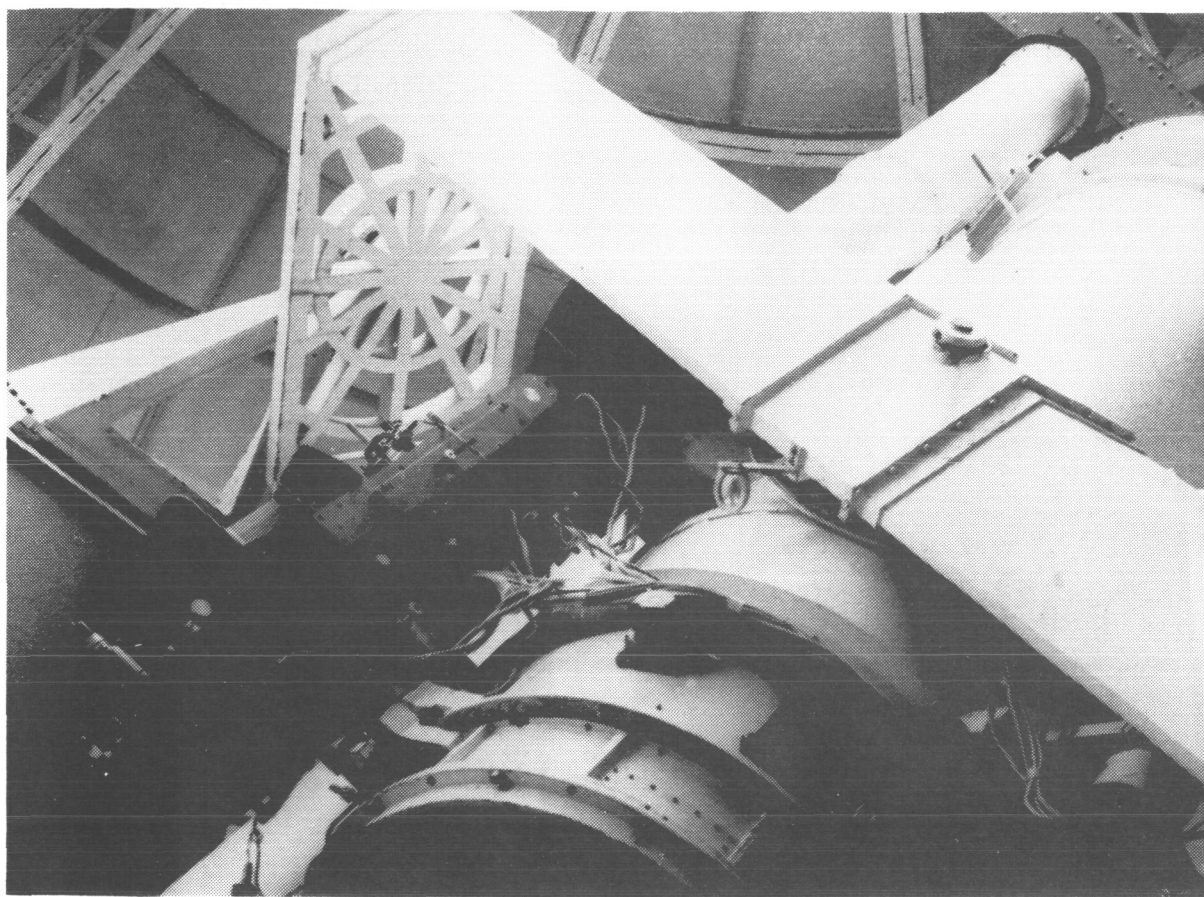


Figure 1. The filter FPSS and its solar refractor 28 cm diameter at Meudon Observatory.

Out best solar photospheric images yet recorded with FPSS centered on the core or the wings of $H\alpha$, D_3 , b_1 or FeI 5576 Å lines have true resolutions of around 0.8 arc sec. (half width of the modulation transfer function peak) while the theoretical diffraction limit of the telescope alone is 0.5 arc sec.

2. Radial Velocity Imaging with FPSS

In order to illustrate the prospects for high angular resolution mapping of the Stokes parameters and magnetic fields at the solar surface, using FPSS with a Stokesmeter modulation device described below, we produce here some results on quantitative imagery of radial velocities obtained at Meudon Observatory.

Doppler shifts are converted into contrast variations by setting the transmission band to isolate the wings of absorption lines such as the strong NaD_1 , or the non magnetic FeI 5576.

2.1 Photospheric Oscillations

Figure 2 shows a direct image of the photospheric 300 sec. oscillations, taken at -0.08\AA from the core of NaD_1 on June 1st, 1981. The areas where matter is moving upward are shown as dark and downward motions appear light. The structures of the oscillating elements are detailed at the solar surface with a resolution of 500 km. A direct Doppler shift calibration produces an average vertical, oscillating maximum velocity of 220 ± 40 m/sec. which has to be slightly increased when a residual image blurring

effect is taken into account. A movie sequence of images indicates a period of 285 ± 25 m/sec. for these coherent oscillations. A power spectrum of the velocity field is generated with a laser source and shown in Figure 3. The bright ring of radius 0.50×10^{-3} rad. results from the concentration of oscillating elements with a size around 4.5 Mm (half the average distance between elements). Another ring of radius 0.2×10^{-3} rad., although blurred with the central peak in the reproduction of Figure 3, characterizes another superimposed coherent oscillation system with a typical size of around 12 Mm. This oscillating system is separated from the first one by optical filtering in Figure 4. An average maximum oscillation velocity can also be extracted from Dopplershift histograms and gives 650 ± 50 m/sec. for this second coherent system (see Dollfus, 1984, Dollfus et al. 1984).

2.2 Motion in sunspot groups

Radial velocity mapping in sunspot areas requires a selection of the contrasts generated by the Doppler shifts from the true intensity features. We practice photographic subtraction of the images taken in the two wings of the line. Figure 5 shows such a radial velocity field map, around a sunspot group on Sept. 16, 1982. A very wide field is preserved by the photographic image subtraction technique. The velocities over the spot areas are detailed in Figure 6, either by photographic image subtraction processing (at left) or by computer processing (at right).

In Figure 7, the radial velocity field is imaged around a sunspot near the limb on Sept. 14, 1982 at a photospheric altitude which, according to the model of NaD_1 line formation, corresponds to the minimum of temperature. The motion is receding at left and approaching at right. It does not correspond to the divergent "Evershed" flow at the photospheric level, nor to the convergent "Abetti" motion at the chromospheric level, but indicates, for this intermediate altitude, a retrograde vortex.

3. Developments Expected to Follow

The results already obtained on quantitative imagery with FPSS indicate that we can record as well the Zeeman $\pm \sigma$ shifts in magnetic sensitive lines to produce high resolution maps of the longitudinal magnetic field. A V Stokes parameter modulator is placed in front of the filter for this purpose and a solid-state bidimensional array detector is expected to make possible real time interactive analyses.

After tests and improvements at the focus of the 28 cm Meudon solar refractor, the whole birefringent filter and peripheral instruments are planned to be moved to the Canaries Islands to be attached at the focus of the THERMIS French solar telescope, 90 cm diameter, specifically designed and dedicated to high resolution solar work. This telescope is scheduled to be operational in the near future at a low turbulence selected site of Teneriffe Island, altitude 2400 meters (see J. Rayrole, this book). Quantitative images of Stokes parameters are expected to be produced with angular resolutions better than 250 Km at the solar surface. A Stokes parameters modulator is needed in front of the filter FPSS. The principle of the device is described below.

4. Double Genova-Cross Stokes Modulator

For high angular resolution imaging, exposure times of at least a fraction of a second are needed, with photographic emulsions as well as with solid-state detectors, and a fast modulating device is not required. To accommodate the whole field of 36 mm diameter, large size birefringent pieces are needed. The modulator must not produce any detectable signal when the incident light beam is not polarized.

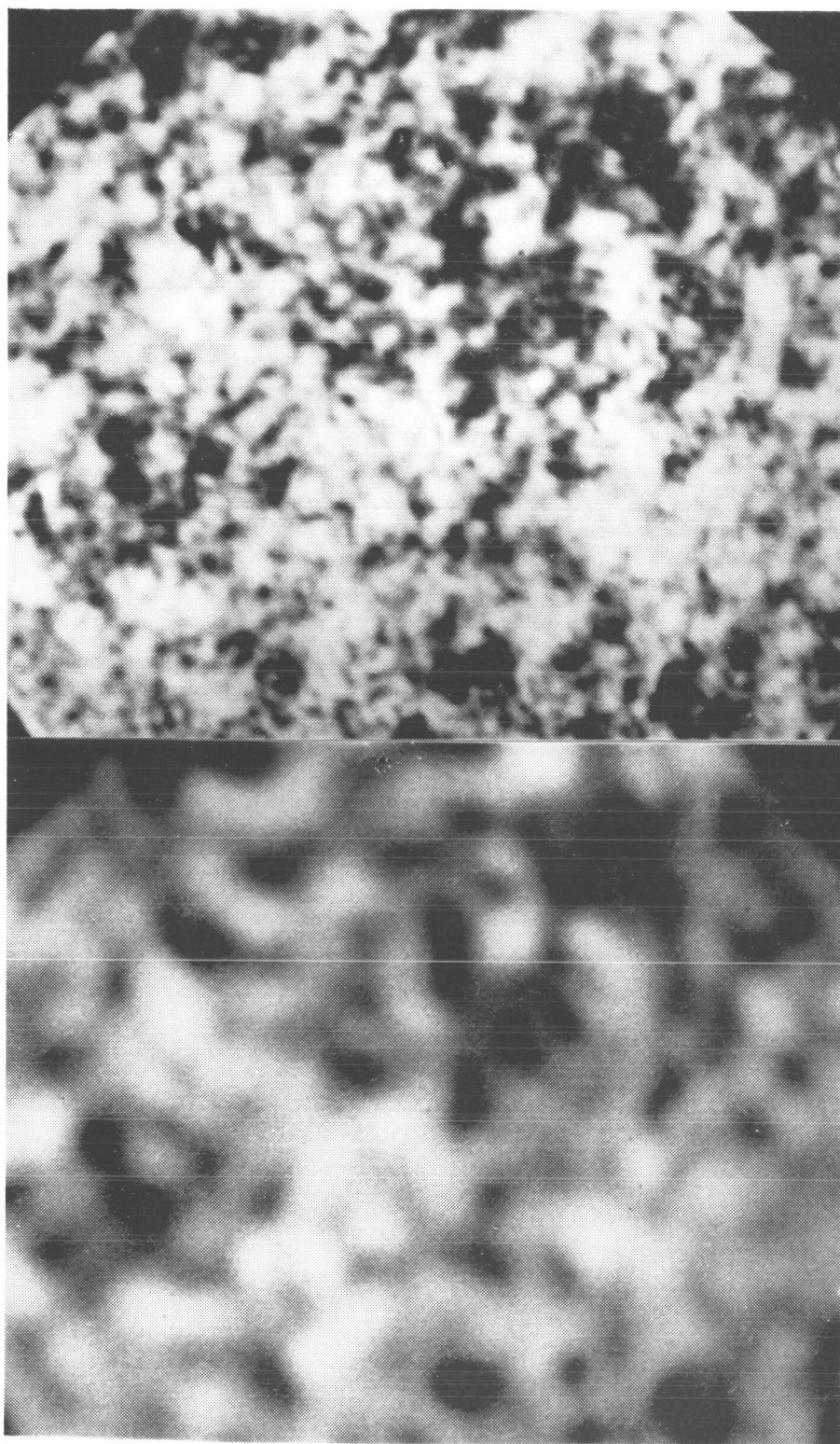


Fig. 2. Direct image of the photospheric oscillation of period 300 sec. FPSS filter in the blue wing of NaD_1 . Field = $145 \times 124 \text{ (Mm)}^2$. Upward motions are light. Downward motions are dark.

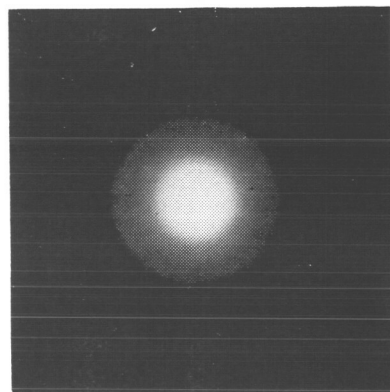


Fig. 3. Power spectrum of the radial velocity field of Figure 2.

Fig. 4. Spatially filtered image of the field of Figure 2, isolating the coherent oscillatory elements of 10.15 Mm average diameter.

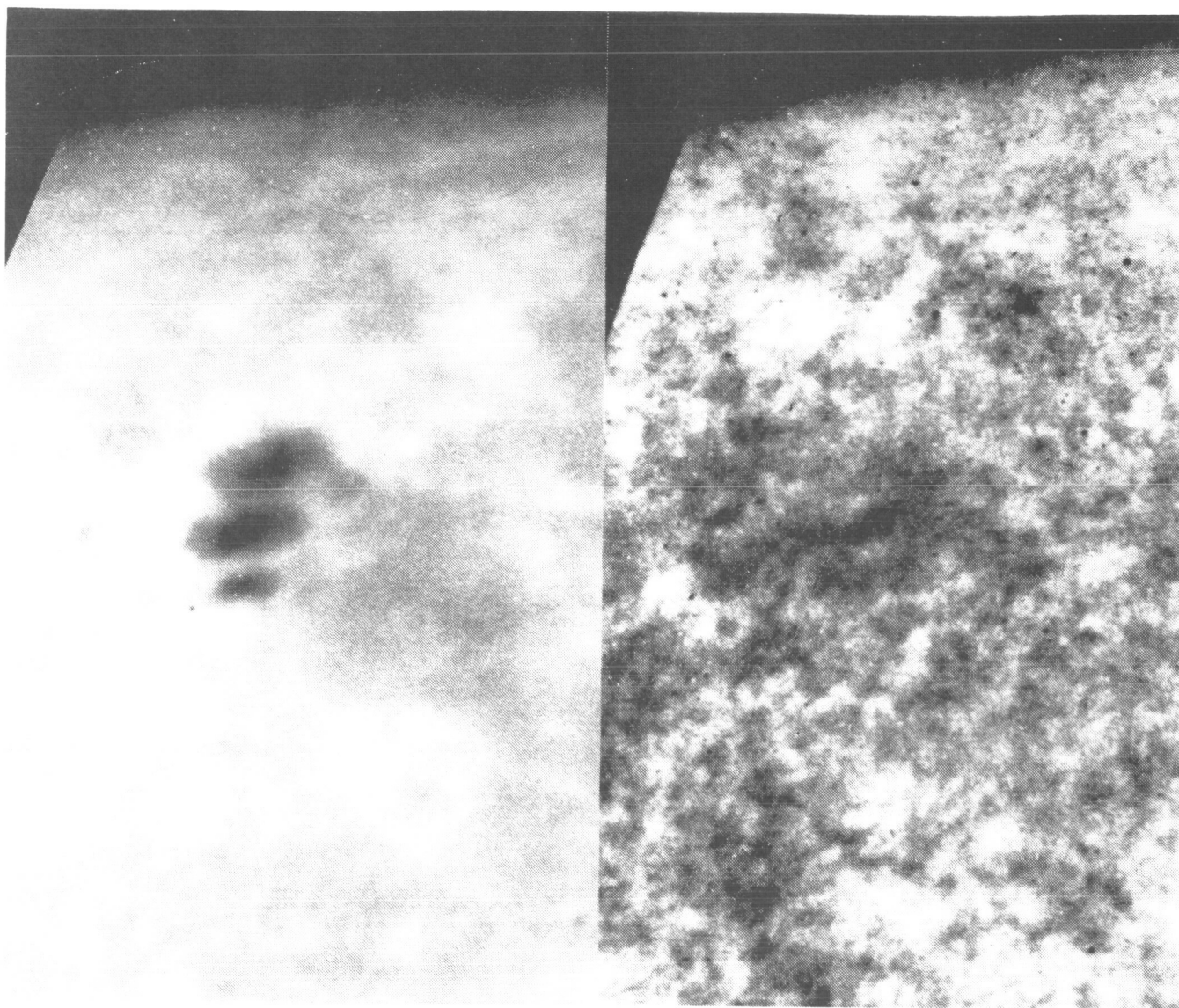


Fig. 5. Radial velocity field image with FPSS, Sept. 16, 1982, 13:14 UT. (Left): Image centered on NaD_1 line (level of minimum temperature). (Right): Velocity field by photographic subtraction of images taken in each wing of NaD_1 . The limb of the Sun is at top. Field = 2.5×4.0 (arc min)².

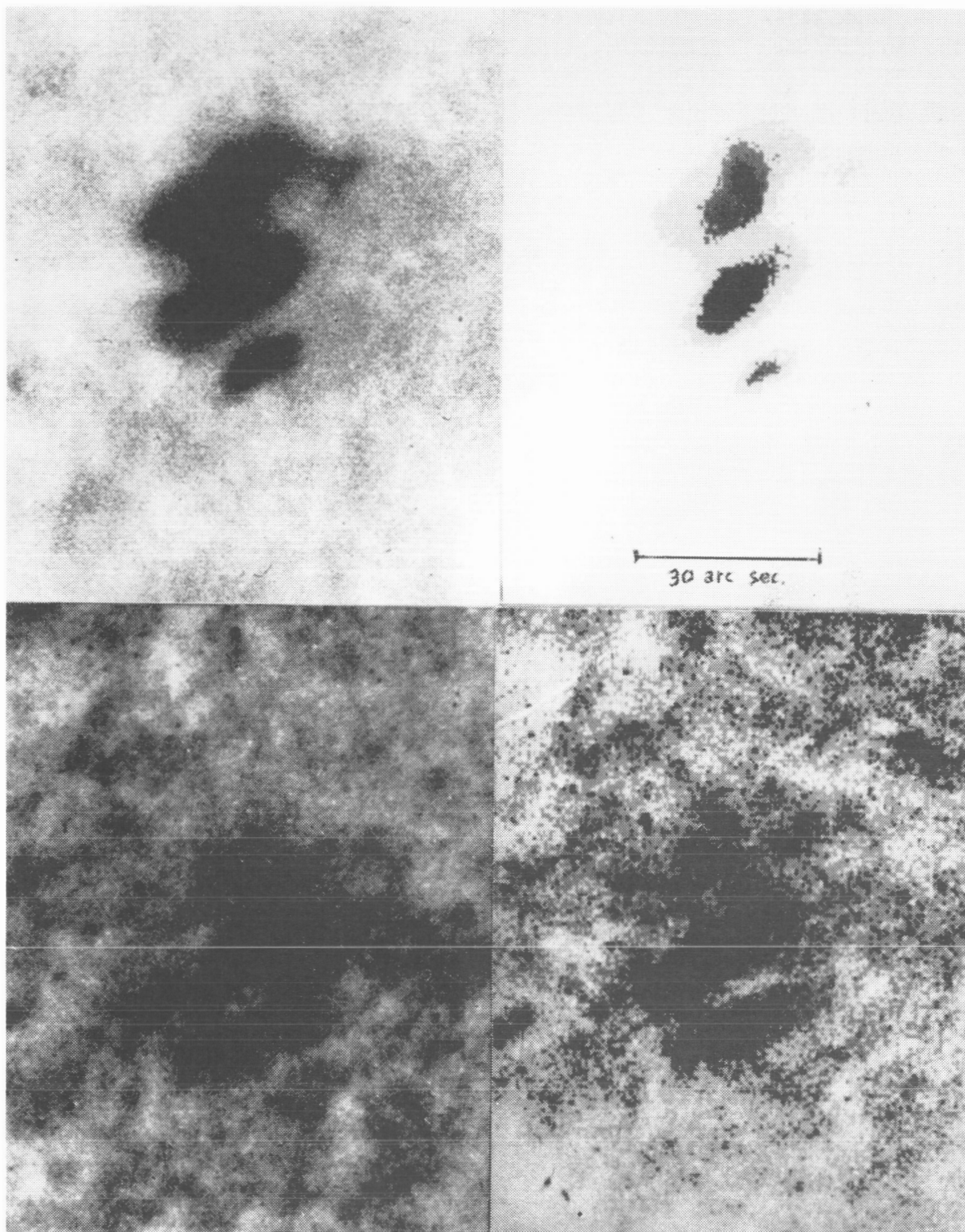


Fig. 6. Radial velocity field around a spot, Sept. 16, 1982, 13:14 UT. (Top left): Image centered on NaD_1 line. (Top right): Isophotes in continuum. (Bottom left): Velocity field (photographic subtraction). (Bottom right): idem (computer subtraction).

The so called “cross talk” – partial mixing of the Stokes parameters – has to be carefully prevented. Such are the major considerations which led to the design of the following device.

4.1 Principle of the Double Rotating, Half Wave Plate Modulator

Optical components are placed in front of the birefringent filter FPSS as shown in Figure 8. The first polarizer of the filter P has a reference position angle of 0° and is associated with a fixed quarter wave plate $(\lambda/4)_0$ at $+45^\circ$. Then two quarter wave plates $(\lambda/4)_1$ and $(\lambda/4)_2$ are added. Their axes are initially set at $\theta_1 = \theta_2 = 0^\circ$ and they can rotate around the light beam axis by steps of exactly 45° . When the front plate $(\lambda/4)_1$ is incremented by $+45^\circ$, the other plate $(\lambda/4)_2$ is stepped by -45° at the same time, but for each even step of the front plate only.

During a complete cycle of 8 increments P_i , the front plate has occupied 8 positions at 45° from 0° to $+360^\circ$, the second plate 4 positions from 0° to -180° and the six Stokes parameters have been obtained as indicated in the Table; the $\pm V$ parameters are produced twice.

Finally, the following image subtractions produce the required Stokes parameter maps.

$$(P_3 - P_7) \rightarrow 2Q$$

$$(P_2 - P_6) \rightarrow 2U$$

$$(P_1 - P_5) \rightarrow 2V$$

Alternatively, the map for $2V$ can be produced by one of the three other combinations $(P_8 - P_4)$, $(P_8 - P_5)$ or $(P_1 - P_4)$.

4.2 Poincare Analysis

The way the Stokes parameters are displayed is readily understood when drawing the retardation paths of the plates on a Poincare sphere in which the 6 Stokes parameters are located as shown in Figure 10. The 100% linearly polarized light at position angle 0° produced by the polarizer P is located in a . Because all the angles are doubled, the active axis of plate $(\lambda/4)_0$ is placed at $(+45^\circ) \times 2$ which is along ob ; its phase retardation of $+45^\circ$ rotates a around ob by 90° to place a at e . Then, for each position P_i , the locations of the two rotating plates' axes on the sphere and the resulting retardation paths are detailed in Table I together with the Stokes parameter that is finally isolated.

4.3 Mueller Matrices Formulation

Assuming a correct retardation for each plate, the whole polarization matrix can be written and each Stokes parameter can be extracted with proper values of θ_1 and θ_2 . For example, for $+Q$ we have:

$$Q_1 = -\sin 2\theta_2 [Q_0 \cos^2 2\theta_1 + V_0 \sin 2\theta_1 \cos 2\theta_1 + V_0 \sin 2\theta_1] \\ + \cos 2\theta_2 [Q_0 \sin 2\theta_1 \cos 2\theta_1 + V_0 \sin^2 2\theta_1 - V_0 \cos 2\theta_1]$$

If the position angles θ_1 or θ_2 are not exactly multiples of 45° but are $\theta_1 = n \lambda/4 + \alpha$ and $\theta_2 = n \lambda/4 + \beta$, the photometric modulation which is produced when measuring Q is:

$$m_Q = \frac{Q_0 + 2 \beta V_0}{1 - 2 \alpha V_0}$$

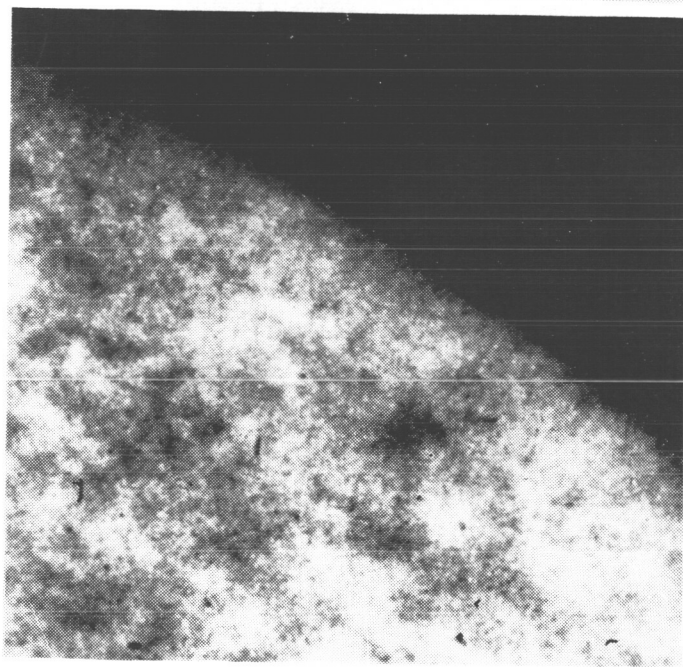
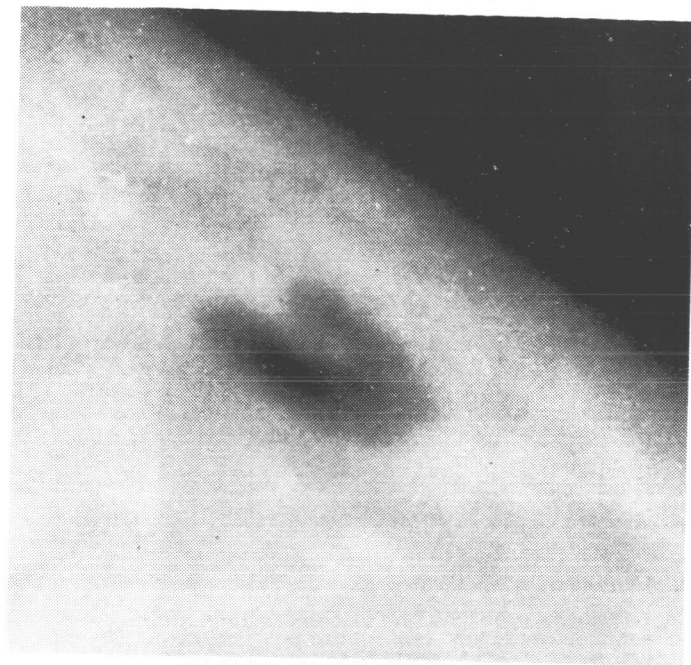


Fig. 7. Sept. 14, 1982 at 10:18,50 UT. Top: Sunspot near limb, filter centered at NaD_1 . Bottom: Radial velocity field. Photographic subtraction of images taken at ± 0.10 Å from NaD_1 center. In dark: approaching. In light: receding.

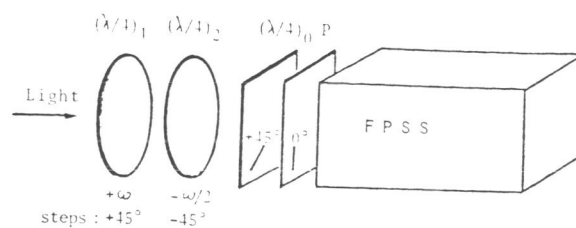


Fig. 8.

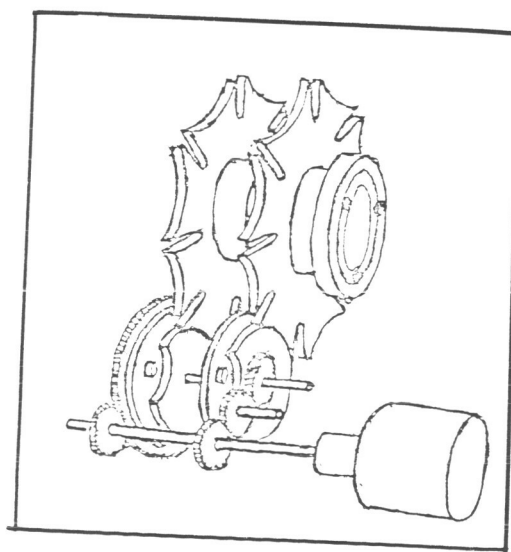


Fig. 9.

Table I.

Position	Position angle of plates axes		Axes on the Poincaré sphere		Retardation paths on Poincaré sphere	Parameter selected
	θ_1 deg.	θ_2 deg.	θ_1	θ_2		
P_1	0	0	oa	oa	af	+V
P_2	+45	0	ob	oa	afb	+U
P_3	+90	-45	oc	od	afa	+Q
P_4	+135	-45	od	od	afae	-V
P_5	+180	-90	oa	oc	afde	-V
P_6	+225	-90	ob	oc	afd	-U
P_7	+270	-135	oc	ob	afc	-Q
P_8	+315	-135	od	ob	afcf	+V

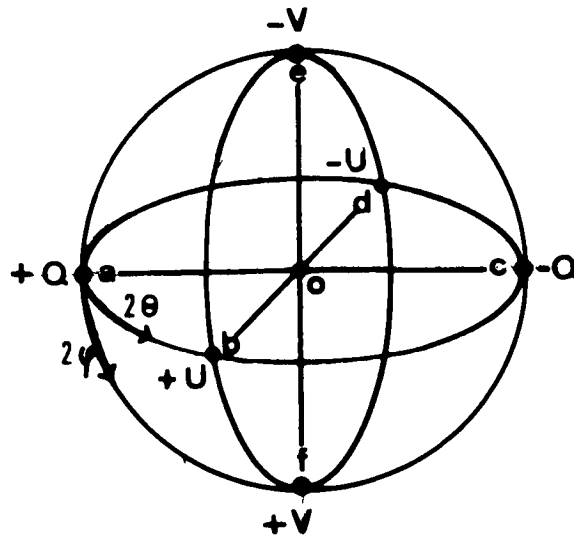


Fig. 10.

U_o is injected in the signal and the resulting "cross talk" could be serious when Q_o is small and U_o large. If $\beta = 1^\circ$, a spurious signal of $3 \times 10^{-2} V_o$ is introduced.

Assuming now correct position angles but phase retardation defects for the plates, we have $\varphi(\lambda) = \pi/2 + \epsilon(\lambda)$, and a similar computation produces:

$$m_Q = \frac{Q_o - \epsilon V_o}{1 - \epsilon V_o}$$

It is V_o which is not injected in the signal. Again, a plate retardation defect of 1° introduces a signal of $1.6 \times 10^{-2} \times V_o$.

4.4 The Retardation Plates

The three quarter wave plates have 36 mm diameter. To secure the retardation tolerance required, these plates are achromatic compounds of the Beckers type, made of a quartz plate cemented with a synthetic magnesium fluoride MgF_2 plate. The ratio between the two thicknesses is 1.16/1. The residual wavelength dependence of the birefringence and other details are found in Beckers (1971).

4.5 Mechanical Design

The two rotating retardation plates are held each in an 8 position Genova-cross mounting, which gives mechanically fixed position angles at exactly 45° with the accuracy required. Proper gears associate the driving shafts of the two Genova-crosses with relative rotation rates of +1 and -1/2. The driving motor operates the gears to produce a step rotation of 45° in less than 100 millisec; then the plate rests at its exact new position until the drive shaft makes a new revolution to initiate a second 45° step. The second plate is operated exactly at the same time in the opposite direction, once every two movements of the first plate.

The principle of the double Genova-cross mechanism is outlined in Figure 9.

Each of the 6 Stokes parameters $\pm Q$, $\pm U$ and $\pm V$ can be individually selected at will, or they can be sequentially driven at a maximum rate of several positions each second.

Acknowledgements: We are particularly indebted to Herve de Larquier for his analysis, tests and laboratory work about the double Genova-cross modulator and its optimization.

References

- Beckers, J. M., 1972, *Applied Optics*, 10, 973-975.
Dollfus, A., 1984, *Proc. Japan-French Seminar*, p. 177 (Pub, CNRS and Observatoire de Paris).
Dollfus, A., Banos, G., Crussaire, D., Jordy, M., Michau, V., Morice T., and Titeaux P., 1984, *Comptes rendus Acad. Sci. Paris*.

THE SAN FERNANDO OBSERVATORY VIDEO STOKES POLARIMETER

PAUL H. RICHTER, LOUIS K. ZELDIN* AND TIMOTHY A. LOFTIN

California State University, Northridge

Abstract. A study has been conducted to determine the suitability of the San Fernando Observatory's 61 cm (24 inch) aperture vacuum solar telescope and 3 m (118 inch) focal length vacuum spectroheliograph for Stokes Polarimetry measurements. The polarization characteristics of these two instruments have been measured by determining their Mueller matrices as a function of telescope orientation, field angle, wavelength, grating type, and position of the measuring beam in the telescope entrance window. In general, the polarizing and depolarizing properties are small so that inversion of the system Mueller matrix will permit the accurate measurement of Stokes profiles for vector magnetic field determination. A proposed polarimeter design based on the use of a TV camera system to simultaneously scan six different polarization components of a given line profile is described. This design, which uses no rotating optics or electronic modulators and makes efficient use of the available irradiance, promises to yield high quality vector magnetograms.

1. Introduction

The San Fernando Observatory has for some 15 years produced high quality solar magnetograms with the observatory's 61 cm (24 inch) aperture vacuum solar telescope and 3 m (118 inch) focal length vacuum spectroheliograph, using a photographic subtraction process to obtain the longitudinal magnetic field component. More recently consideration has been given to using this system in conjunction with a new polarization analysis-video detection system to produce vector magnetograms by means of Stokes polarimetry. The proposed analysis and detection scheme is based on the use of a combination of non-polarizing and polarizing beamsplitters arranged so as to produce six images of the spectroheliograph exit slit, each representing different components of polarization. These six images are juxtaposed on the face of a TV camera system so that the horizontal scan samples the resulting six line profiles sequentially, while the vertical scan permits sequential sampling of the profiles corresponding to the different points of the solar image falling on the entrance slit of the spectroheliograph. The video signal from the camera system is fed to a video cassette recorder (VCR) for future analysis carried out by means of an image digitizing system. This analysis and data reduction consists of signal integration, photometric correction, polarization correction and algebraic manipulation to extract the Stokes profiles $S(\lambda)$ for each image point scanned. Mechanical scanning of the spectroheliograph

* Currently at Rockwell International, Rocketdyne Div., Canoga Park, CA. U.S.A.

moves the entrance slit across the solar image sweeping out a rectangular area, each line of which is treated as described above, so that the result is a complete Stokes profile map for the area scanned.

The following sections of the paper describe the optical and electronic systems in more detail, discuss the data reduction problem, and present measurements of the polarizing characteristics of the telescope and spectroheliograph, the latter illustrating the suitability of this system for obtaining polarimetric data.

2. Polarimeter Design

The design of the proposed Video Stokes Polarimeter has been driven by a desire to avoid problems associated with rotating polarizers and/or retardation plates and electro-optic devices such as have been used in various modulating schemes, and to take maximum advantage of the intensity and resolution provided by the telescope and spectroheliograph through the combined use of efficient beam division and image multiplexing.

The method of analysis can be understood by referring to Figure 1, which shows a block diagram of the optical system. The beam incident on the telescope is characterized by the Stokes vector

$$S = \{ I, Q, U, V \},$$

where each element is a function of position on the solar image, and wavelength, i.e., $S = S(x, y, \lambda)$, where x and y are coordinates in the telescope image plane and λ is the wavelength associated with a given point in the absorption line profile under study.

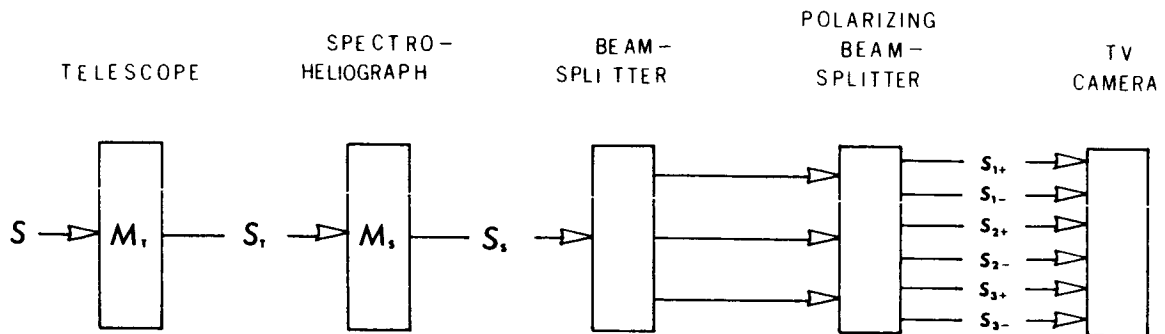


Figure 1: Block diagram of the Video Stokes Polarimeter.

In passing through the telescope the beam is altered by the polarizing characteristics of the instrument so that the telescope image is characterized by the Stokes vector

$$S_T = M_T S \quad 1)$$

where the telescope properties are represented by the Mueller matrix M_T . The elements of this matrix are functions of the mean wavelength $\bar{\lambda}$ of the line under study, position in the field of view, and the telescope hour angle and declination, i.e., $M_T = M_T(x, y, \bar{\lambda}, h, \delta)$.

These functional dependences stem from the reflecting optical design of the telescope and will be discussed in greater detail in the section dealing with the measurement of M_T .

On emerging from the telescope the beam passes through the spectroheliograph and suffers additional change so that the image formed at the exit slit of the instrument is characterized by the Stokes vector

$$S_S = M_S S_T. \quad 2)$$

The spectroheliograph matrix M_S is a function of position y along the exit slit, mean wavelength λ , and the particular grating selected, G . Details are discussed in the section dealing with the measurement of M_S .

The polarimeter optics consists of a 1:1 imaging system containing a non-polarizing beamsplitter and a polarizing beamsplitter assembly. The former merely divides the beam into three equal parts, while the latter performs an analysis of each of these three beams into orthogonal polarization states. The result is six images of the spectroheliograph exit slit, each corresponding to a different state, and these are juxtaposed on the TV camera tube faceplate.

The analysis section (polarizing beamsplitter assembly) consists of three identical Wollaston prisms preceded by three different linear retarders, the first having zero retardance (glass compensating plate), the second π retardance with fast axis oriented $22\frac{1}{2}^\circ$ relative to one of the transmission axes of the Wollaston prisms, and the third $\pi/2$ retardance with fast axis oriented 45° relative to the Wollaston transmission axis.

The Mueller matrices for these devices are given by the following:

$$\text{WOLLASTON PRISM} \quad M_{\pm} = \frac{1}{2} \begin{bmatrix} 1 & \pm 1 & 0 & 0 \\ \pm 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad 3a)$$

$$\text{HALFWAVE PLATE} \quad M_H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad 3b)$$

$$\text{QUARTERWAVE PLATE} \quad M_Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad 3c)$$

Thus, the three sections of the analyzer have the following matrices:

$$M_{\pm} = \frac{1}{2} \begin{bmatrix} 1 & \pm 1 & 0 & 0 \\ \pm 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad 4a)$$

$$M_{\pm} M_H = \frac{1}{2} \begin{bmatrix} 1 & 0 & \pm 1 & 0 \\ \pm 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad 4b)$$

$$M_{\pm} M_Q = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & \pm 1 \\ \pm 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad 4c)$$

and these operate on the incident beams to produce the following outputs:

$$S_{1\pm} = \frac{1}{6}(I_S \pm Q_S) \begin{bmatrix} 1 \\ \pm 1 \\ 0 \\ 0 \end{bmatrix}, \quad S_{2\pm} = \frac{1}{6}(I_S \pm U_S) \begin{bmatrix} 1 \\ 0 \\ \pm 1 \\ 0 \end{bmatrix}, \quad S_{3\pm} = \frac{1}{6}(I_S \pm V_S) \begin{bmatrix} 1 \\ 0 \\ 0 \\ \pm 1 \end{bmatrix}. \quad 5a-c)$$

The corresponding beam irradiances are,

$$I_{1\pm} = \frac{1}{6}(I_S \pm Q_S), \quad I_{2\pm} = \frac{1}{6}(I_S \pm U_S), \quad I_{3\pm} = \frac{1}{6}(I_S \pm V_S), \quad 6a-c)$$

so that sums and differences of adjacent beam irradiances have the values,

$$\sum_1 = I_{1+} + I_{1-} = \frac{I_S}{3}, \quad \sum_2 = I_{2+} + I_{2-} = \frac{I_S}{3}, \quad \sum_3 = I_{3+} + I_{3-} = \frac{I_S}{3}, \quad 7a-c)$$

$$\Delta_1 = I_{1+} - I_{1-} = \frac{Q_S}{3}, \quad \Delta_2 = I_{2+} - I_{2-} = \frac{U_S}{3}, \quad \Delta_3 = I_{3+} - I_{3-} = \frac{V_S}{3}, \quad 8a-c)$$

thus permitting a determination of the Stokes vector $S_S = \{I_S, Q_S, U_S, V_S\}$.

Combining equation 1) and 2) we have,

$$S_S = MS, \quad 9)$$

where $M = M_S M_T$ is the overall system matrix. Thus, by inverting the latter we obtain

$$S = M^{-1} S_S, \quad 10)$$

and substitution of the measured elements given by equations 7) and 8) into the above yields the input Stokes vector S .

The above discussion presents the essential principles involved in the operation of the polarimeter. In practice, a number of factors will result in a somewhat more complex situation than that illustrated, thereby complicating the algebra but otherwise not affecting the general result. Thus for example, while a retardation plate will produce a given retardance at a given wavelength, even an "achromatic" design cannot provide a constant value of this parameter over a finite range of wavelengths. Furthermore, unless measures are taken to eliminate multiple passes of the beam through the plate as a result of reflection from various surfaces in the system, the retardance and polarizing properties of the analyzer will become strong functions of wavelength (Holmes, 1964; Smith, 1969). Also, the "non-polarizing" beamsplitter and the reflective imaging optics can be expected to introduce some polarization.

These effects can all be included in the analysis, and as long as the Mueller matrix for each element is known, or can be measured, the incident Stokes vector S can be found by procedures similar to that described above. Calibration procedures are discussed in a subsequent section, but before coming to these we first describe the implementation of the proposed measurement scheme and the data acquisition and reduction methods to be employed.

3. Optical System

The telescope and spectroheliograph optical systems are shown in Figures 2 and 3, and the proposed polarimeter optics in Figure 4. The functioning of the latter may be understood by noting that the telescope field lens, located immediately above the coudé focus, images the telescope entrance pupil (entrance window) at the spectroheliograph diffraction grating, thus forming the exit pupil for the instrument. Mirror M1 in the polarimeter (Figure 4) serves two functions. First, it collimates the beam for refocusing by mirror M3, thus providing a compact, folded aberration compensated 1:1 imaging system (Czerny-Turner design). Second, it images the spectroheliograph exit pupil at mirror M3, which also serves the dual purposes of imaging and splitting the beam into three equal irradiance parts, the mirror itself being a sectioned sphere with independent articulation for each section. The polarizing beamsplitter assembly is mounted immediately in front of the focal plane located at the TV camera tube fiber-optic faceplate. Figure 5 shows a detail of the polarizing beamsplitter and the image format on the faceplate, with typical dimensions corresponding to the use of an echelle grating providing a linear dispersion of $2 \text{ mm}/\text{\AA}$, and a slit length of 15 mm corresponding to 4.23 arc minutes across the solar disc.

4. Data Acquisition and Reduction

The TV camera converts the faceplate irradiance distribution into a video signal whose voltage $v(t)$ is proportional to the irradiance at the point scanned by the

SOLAR TELESCOPE OPTICAL SYSTEM

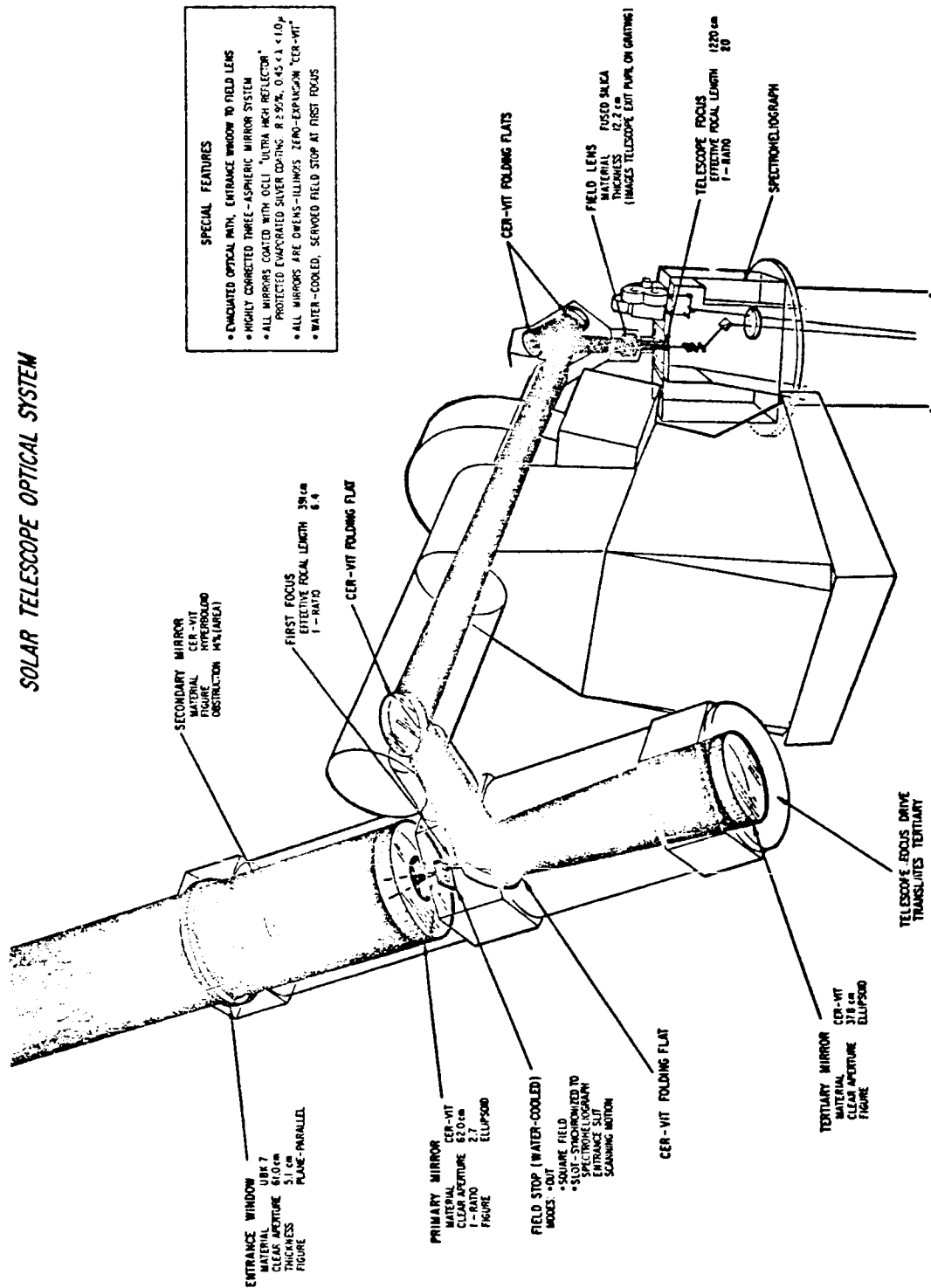


Figure 2: Telescope optical system.

SPECTROHELIOGRAPH OPTICAL SYSTEM

SOLAR IMAGE
IMAGE SCALE 50x/100x
F-RATIO 20

TELESCOPE FIELD LENS
IMAGES TELESCOPE EXIT PUPIL THRU LUTTROW LENS ONTO GRATING

INTERFERENCE PREFILTER

TO 100mm FILM AND TWIN FILM TRANSPORT CAMERA
A. FRAME SIZE: 56 x 56 mm
B. PRECISION PIN REGISTRATION FOR SUBSEQUENT PHOTOGRAPHIC SUBTRACTION

MATCHED PAIR OF EXIT SLITS
LENGTH 57 mm
RADIUS OF CURVATURE 2200 mm

COMBINED TILT PLATE AND NEUTRAL DENSITY FILTER
A. OBJECT'S IMAGE AT EXIT SLIT
B. BALANCES TWO EXIT BEAM INTENSITIES

INTERCHANGEABLE OPTICS
A. SCANNED ROLL OF 100mm FILM
B. POLARIZED EXIT BEAMS, MAGNETIC FIELD RECORDING MODE
C. EQUAL-SPLIT BEAM SPLITTER-COMPLER RECORDING MODE
D. THICK PLATE - SINGLE CHANNEL RECORDING MODE

BEAM SPLITTER

SPECTRUM LINE POSITIONING SERVO
INCORPORATES SELECTING SLIT AND BEAM ORDER STIMULATES LIGHT INTO 'RED' AND 'BLUE' CONTRIBUTIONS FROM SPECTRUM LINE

LIGHT CHOPPER
ALTERNATING CHOPS 'RED' AND 'BLUE' BEAMS

TO PHOTOMULTIPLIER

VACUUM TANK

LUTTROW TRIPLE LENS
SLITLESS TELEPHOTO

PLANE-REFLECTANCE GRATING NO. 1
GROOVES/MM 600
BEVELED WIDTH 254 mm
GROOVE LENGTH 154 mm
PLATE, 1st ORDER 135 μ
FUSED SILICA BLANK

PLANE-PARALLEL VACUUM WINDOW

SPECTROHELIOGRAPH ROTATION AXIS
270° RATE COMPENSATES FOR DIURNAL ROTATION OF IMAGE AT CLUSE FOCUS

EVACUATED OPTICAL PATH
WINDOW TO GRATING

3-GRATING TURRET

SPECTROHELIOGRAPH SCAN MOTION
LOCATED AT POSITION TELESCOPE EXIT PUPIL
WOULD OCCUPY IN THE ABSENCE OF THE SPECTROHELIOGRAPH OPTICS

ENTRANCE SLIT
LENGTH 57 mm
RADIUS OF CURVATURE 2200 mm
REFLECTING, INCLUDED FLAT JAMS

SPECTRUM LINE CURVATURE COMPENSATING OPTICS
PERFORMS CONTINUOUS ADJUSTMENT OF CURVATURE COMPENSATION

1/4 PLATE
USED IN MAGNETIC FIELD RECORDING MODE

BARREL SHUTTER

Figure 3: Spectroheliograph optical system.

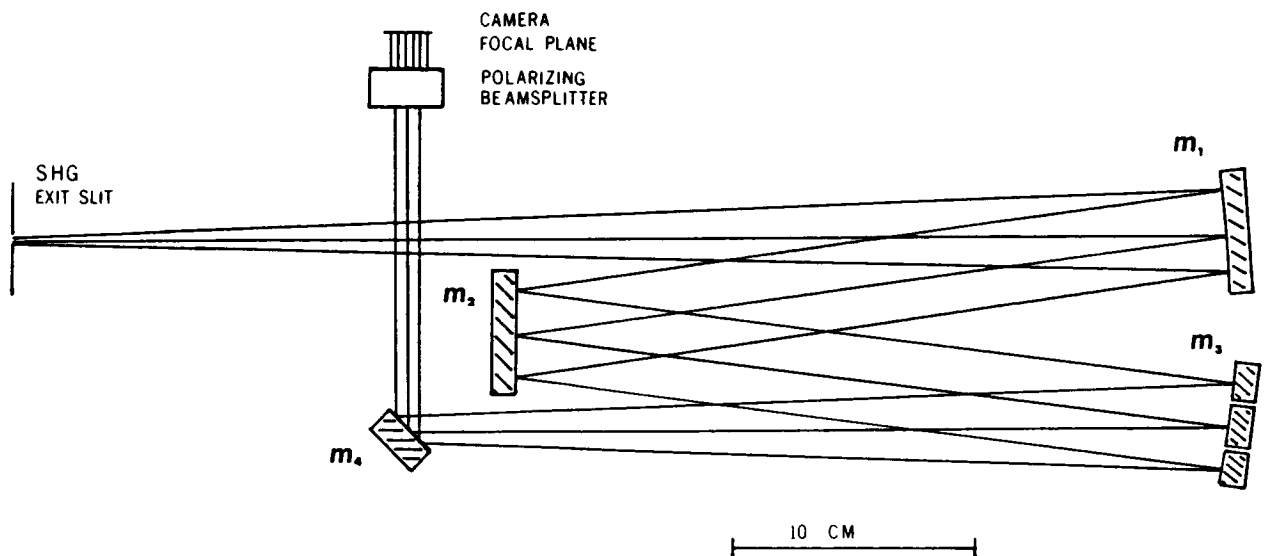


Figure 4: Polarimeter optical system. Mirror M1 collimates the beam and images the spectroheliograph exit pupil at mirror M3, which serves as a non-polarizing beamsplitter and focusing element in a 1:1 Czerny-Turner arrangement.

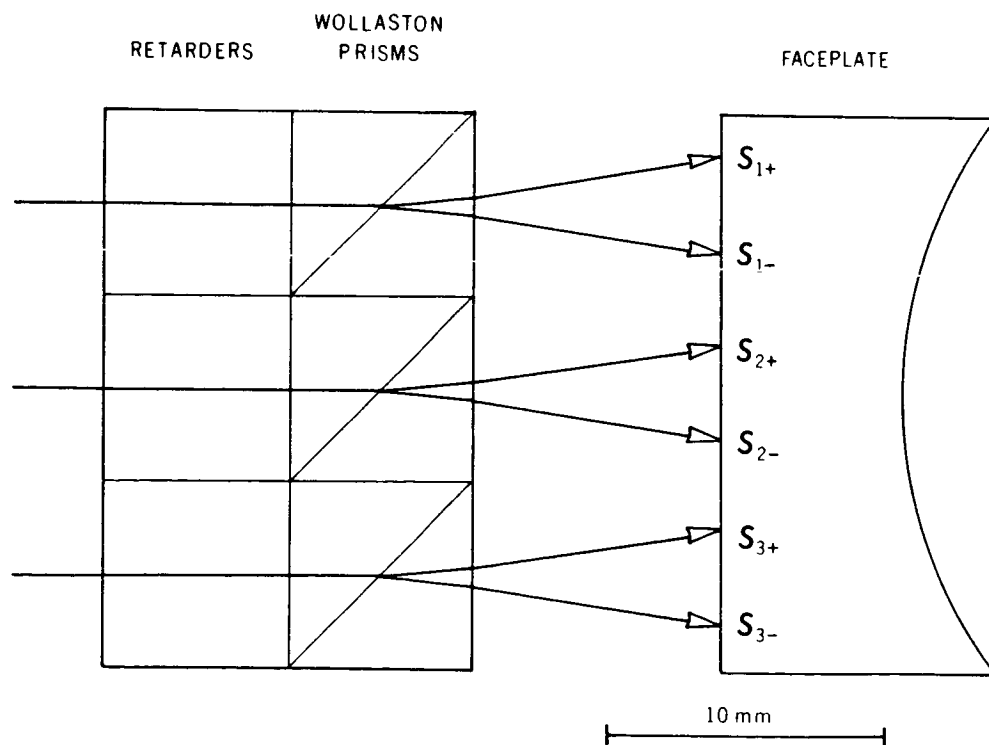


Figure 5: Polarizing beamsplitter assembly.

camera's electron beam at time t . Thus, a typical portion of this signal corresponding to a given horizontal scan line, and hence to a given solar image point, consists of an analog sequence of the six Stokes profiles given by equations 6a-c of section 2. This signal however, suffers from amplitude error resulting from the non-uniform gain characteristics of the camera system, and applying a point by point photometric correction for this error is the first step in the data reduction process.

Since this correction, as well as subsequent ones to be discussed, cannot be carried out with the required precision in real time, i.e., at the standard video rate of 30 TV frames/sec., the raw video signal is first stored on magnetic tape using a commercial VCR (GE RECORDER (DECK) MOD.ICVD 4020X). Subsequent processing is then carried out by transferring TV frames from tape to a digital image processing system (QUANTEX DIGITAL IMAGE MEMORY PROCESSOR, MODEL DS-12) capable of integrating each pixel up to a maximum of 12-bits with a frame format of 256 x 256 pixels. This frame integration can be carried out at video rates so that up to 16 frames can be summed in a little over a half second.

The photometric correction tables and system matrix elements are stored on disc and read out onto a computer RAM in blocks to permit line by line photometric corrections and matrix multiplications to be carried out on the sums and differences represented by equations 7 and 8, the raw data being read from the image processor through an IEEE bus.

The results of these computations are the Stokes profiles $S(x,y,\lambda)$, and since each solar image point (x,y) contains $256/6 \approx 42$ pixels (words) of profile data, and a square sample of the sun 4.23 arc minutes on a side contains $256 \times 256 = 65536$ pixels, a single Stokes profile map for such a region involves a considerable amount of data. In fact, since a full vector magnetic field map of the region would require only three pixels (words) per solar image point, there is something like a factor of 14 redundancy in such a representation, and it would be desirable to eliminate this prior to final storage of the results.

An approach to this data reduction problem which has the additional merit of simultaneously improving the overall signal to noise ratio, is to compute weighted integrals of the Stokes profiles for each solar image point, in terms of which the three magnetic field components can be directly calculated and displayed.

As an example of the method we consider the Seares formulae for the Stokes profiles in terms of the line profile $p(\lambda)$, Zeeman splitting $\Delta\lambda$, inclination of the magnetic field vector γ , and transverse azimuth angle α ,

$$I(\lambda) = \frac{1}{4}(1 + \cos^2 \gamma) [p(\lambda + \Delta\lambda) + p(\lambda - \Delta\lambda)] + \frac{1}{2} \sin^2 \gamma p(\lambda), \quad 11)$$

$$Q(\lambda) = \frac{1}{4} \sin^2 \gamma \cos 2\alpha [2p(\lambda) - p(\lambda + \Delta\lambda) - p(\lambda - \Delta\lambda)], \quad 12)$$

$$U(\lambda) = \frac{1}{4} \sin^2 \gamma \sin 2\alpha [2p(\lambda) - p(\lambda + \Delta\lambda) - p(\lambda - \Delta\lambda)], \quad 13)$$

$$V(\lambda) = \frac{1}{2} \cos \gamma [p(\lambda + \Delta\lambda) - p(\lambda - \Delta\lambda)]. \quad 14)$$

In the analysis of solar magnetic fields, these formulae are of course of limited validity, but they nevertheless serve to illustrate the ideas involved in the data reduction scheme, and suggest how these ideas might be applied to a more complete field description based on the Stokes profiles.

If we consider an isolated line profile $p(\lambda)$, then by integrating equations 11-14 over all wavelengths we obtain.

$$\int_{-\infty}^{\infty} I(\lambda) d\lambda = \int_{-\infty}^{\infty} p(\lambda) d\lambda = 1, \quad (15)$$

$$\int_{-\infty}^{\infty} Q(\lambda) d\lambda = \int_{-\infty}^{\infty} U(\lambda) d\lambda = \int_{-\infty}^{\infty} V(\lambda) d\lambda = 0. \quad (16)$$

Multiplying equations 12 and 13 by λ^2 , equation 14 by λ , and integrating we obtain

$$\int_{-\infty}^{\infty} \lambda^2 Q(\lambda) d\lambda = -\frac{1}{2} \Delta \lambda^2 \sin^2 \gamma \cos 2\alpha, \quad (17)$$

$$\int_{-\infty}^{\infty} \lambda^2 U(\lambda) d\lambda = -\frac{1}{2} \Delta \lambda^2 \sin^2 \gamma \sin 2\alpha, \quad (18)$$

$$\int_{-\infty}^{\infty} \lambda V(\lambda) d\lambda = -1 \Delta \lambda \cos \gamma. \quad (19)$$

These three equations may be solved for the Zeeman splitting $\Delta\lambda$, inclination γ , and azimuth α , so that if the indicated integrals are computed from the measured Stokes profiles the magnetic field components can immediately be found and stored. An additional advantage of this method is that the integrals of equations 17-19 are insensitive to line profile shifts caused by Doppler effect or misalignment in the spectroheliograph exit slit, so long as the latter is several linewidths wide and the line being used does not display large damping wings. Also, the effect of finite camera resolution is minimized when these integrals are used to compute the field components.

Because of its simplicity and directness, this approach based on the Seares formulae, in addition to providing a framework for data reduction based on more complete theories, also provides a means for making approximate calculations to permit a "quick-look" capability. In this case, the computation could be done in analog form in real, or near-real time using analog multipliers, integrators and sum and difference amplifiers. Estimates of the performance of such a quick-look system indicate that if the system is photon noise limited, a 100 x 100 pixel vector field map having a limiting sensitivity of 100 gauss could be produced in about 5 minutes, assuming the use of the $\lambda = 6173 \text{ \AA}$ line.

As a final comment it should be mentioned that irrespective of what approach is taken in the final data reduction, the raw video signal stored on cassette can

always be referred to for different kinds of analysis, since it contains all of the Stokes profile data in analog form, and the corrections and data reduction can be conducted at any time, providing only that such basic data as time and telescope coordinates are simultaneously recorded. Also, while the present VCR is adequate with respect to noise level and dynamic range for the storage of video data originating from a TV camera, a recording system possessing lower noise and increased dynamic range would be required if a diode array were substituted.

5. System Calibration-Determination of the System Mueller Matrix

Since the optical design of the telescope and spectroheliograph involves oblique reflection from coated surfaces it has been clear from the outset that corrections must be made for the polarization induced by these instruments, and the method for accomplishing this has been discussed in the preceding section. Here, we address the calibration procedures to be employed that will permit the determination of the correction matrices, and describe measurements already made in an effort to ascertain the severity of these corrections.

The latter consideration is of concern because not only are the corrections time dependent as a result of changes in angle of incidence accompanying solar tracking with a coudé system, but one can expect some depolarization to occur due to inhomogeneous strain induced birefringence in the telescope entrance window, which is subjected to a pressure difference of one atmosphere. If either polarization or depolarization is excessive, inversion of the system matrix will lead to large errors, and the system will be unable to produce the accuracy required for transverse magnetic field measurements.

Thus, in contemplating the use of the telescope-spectroheliograph combination for polarimetric measurements, one of the first steps undertaken was a systematic determination of the telescope and spectroheliograph Mueller matrices, M_T and M_S , under a representative range of parameter values.

A. Measurement of M_T and M_S .

The method used to determine these matrices is based on the technique of ellipsometry, in which the polarization change suffered by an incident linearly polarized beam as it passes through a system is determined by nulling out the elliptically polarized beam that emerges, using a retardation plate and linear polarizer combination (Bashara, 1969). It possesses the great virtue of simplicity and can produce high accuracy using only the human eye as a detector.

As the details and results of these measurements have been published elsewhere (Zeldin, 1981), we will here present only an outline of the method and a summary of the results.

The measurement system is shown schematically in Figure 6.

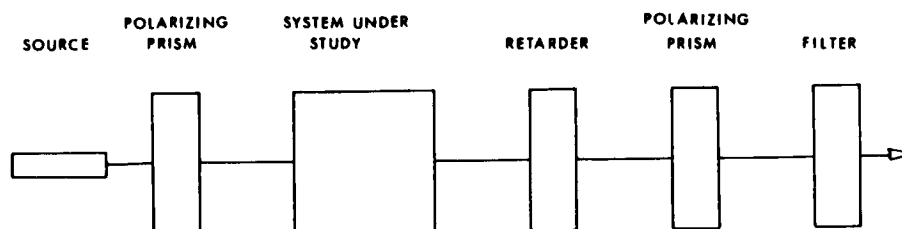


Figure 6: Block diagram for ellipsometric measurement of system Mueller matrices.

Polarizer P produces a linearly polarized beam whose azimuth relative to the system reference axis is p , and whose Jones vector is consequently

$$\begin{bmatrix} J_x \\ J_y \end{bmatrix} = \begin{bmatrix} \cos p \\ \sin p \end{bmatrix} \quad (20)$$

On passing through the system the state of polarization of this beam is modified and is represented by the Jones vector

$$\begin{bmatrix} J'_x \\ J'_y \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} J_x \\ J_y \end{bmatrix} \quad (21)$$

where the Jones matrix $H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$ represents

the polarizing properties of the system being measured.

The emergent beam will in general be elliptically polarized, and by adjusting the azimuth setting of the retarder R it is possible to convert this beam to a linearly polarized state and subsequently null this by adjustment of the azimuth of the analyzing polarizer A. In fact, two such independent null conditions can always be found, and one can show that in terms of the resultant retarder and analyzer azimuths (c, a) , and (c', a') for these nulls, the Jones vector for the emergent beam can be calculated from the following relations (Zeldin, 1981):

$$Z = \frac{J'_x}{J'_y} = \cot p \cot \psi e^{-i\Delta}, \quad (22)$$

$$\tan \psi = \left(\frac{\tan a \tan c' - \tan a' \tan c}{\tan a' \tan c' - \tan a \tan c} \right)^{\frac{1}{2}}, \quad (23)$$

$$2 \tan \psi \cos \Delta = \frac{(\tan a' - \tan a)(1 - \tan c \tan c')}{\tan a' \tan c' - \tan a \tan c}. \quad (24)$$

The elements of the system Jones matrix H may be computed by carrying out the above measurement and computational procedure for three different polarizer azimuths p_1, p_2, p_3 , resulting in the determination of three different z -values, z_1, z_2, z_3 . Then, substituting equation 20 into equation 21 and carrying out the matrix multiplication one obtains the following three linear algebraic equations,

$$H_{12} - Z_1 \cot p_1 H_{21} - Z_1 H_{22} = -\cot p_1, \quad (25a)$$

$$H_{12} - Z_2 \cot p_2 H_{21} - Z_2 H_{22} = -\cot p_2, \quad (25b)$$

$$H_{12} - Z_3 \cot p_3 H_{21} - Z_3 H_{22} = -\cot p_3, \quad (25c)$$

whose solution is given by

$$H_{12} = -\frac{1}{\Delta} \begin{vmatrix} \cot p_1 & Z_1 \cot p_1 & Z_1 \\ \cot p_2 & Z_2 \cot p_2 & Z_2 \\ \cot p_3 & Z_3 \cot p_3 & Z_3 \end{vmatrix}, \quad 26a)$$

$$H_{21} = \frac{1}{\Delta} \begin{vmatrix} 1 & \cot p_1 & Z_1 \\ 1 & \cot p_2 & Z_2 \\ 1 & \cot p_3 & Z_3 \end{vmatrix}, \quad 26b)$$

$$H_{22} = \frac{1}{\Delta} \begin{vmatrix} 1 & Z_1 \cot p_1 & \cot p_1 \\ 1 & Z_2 \cot p_2 & \cot p_2 \\ 1 & Z_3 \cot p_3 & \cot p_3 \end{vmatrix}, \quad 26c)$$

where

$$\Delta = \begin{vmatrix} 1 & Z_1 \cot p_1 & Z_1 \\ 1 & Z_2 \cot p_2 & Z_2 \\ 1 & Z_3 \cot p_3 & Z_3 \end{vmatrix}. \quad 26d)$$

Here, we have considered a normalized Jones matrix for which $H_{11} = 1$, a choice made possible by our lack of interest in absolute amplitude and phase factors in such measurements.

From these results one can compute the corresponding Mueller matrices for the system by well known formulae connecting the two different representations (Gerrard and Burch, 1975). This transformation is possible of course, only when one is dealing with a non-depolarizing system, and it is implicit in the above derivation that this is the case.

In the actual measurements this was also the case since the aperture of the measuring system was small relative to the aperture of the system under study (1 cm vs 61 cm), and the quality of the nulls as well as the accuracy achieved in the measurements attest to the validity of this assumption.

In order to determine the homogeneity of the telescope entrance window, the element that would be most likely to depolarize a full aperture beam coming from the sun, a series of matrices was measured corresponding to different positions of the measuring beam across the window.

In all, a total of 43 telescope matrices M_T , and 18 spectroheliograph matrices M_S were determined in the above manner by varying the entrance window position, field position, wavelength, hour angle and declination, and type of grating. The detailed results of these measurements are available elsewhere (Zeldin, 1981), and have been quite gratifying in that, with the exception of the strong linear polarizing tendency found for the gratings at certain wavelengths, the telescope and spectroheliograph display small polarizing effects, and the inhomogeneity displayed by the telescope entrance window is also quite small, suggesting a correspondingly small depolarizing effect on a full aperture beam.

Tables 1-3 display the Mueller matrices for three representative cases. Here it

M11= .871
M12= -.002
M13= .003
M14= -.129
M21= .005
M22= .861
M23= -.003
M24= -.047
M31= -.007
M32= .006
M33= .859
M34= .070
M41= -.128
M42= .046
M43= -.070
M44= .867

M11= .930
M12= -.061
M13= .001
M14= .035
M21= -.061
M22= .929
M23= -.027
M24= -.011
M31= .001
M32= -.027
M33= -.927
M34= .006
M41= -.035
M42= -.008
M43= -.006
M44= -.928

M11= .524
M12= -.430
M13= .189
M14= .080
M21= -.430
M22= .454
M23= -.163
M24= .007
M31= -.039
M32= -.040
M33= -.215
M34= .034
M41= -.201
M42= -.157
M43= -.102
M44= -.230

TABLE 1: Typical M_T

TABLE 2: $M_S(\lambda = 6309\text{\AA})$

TABLE 3: $M_S(\lambda = 5885\text{\AA})$

can be seen that the magnitude of the diagonal elements of M_T is close to 1.0 while that for the off-diagonal terms is of the order of 0.1 or less. The same is true for the elements of M_S at $\lambda = 6309 \text{ \AA}$ with an echelle grating in position. However, when the wavelength is shifted to $\lambda = 5885 \text{ \AA}$ this same grating displays very strong linear polarization.

Figure 7 illustrates the variation of individual elements of M_T as a function of telescope hour angle and position in the entrance window.

The matrices displayed here have been normalized such that the maximum transmittance for the system is 1.0, i.e., the eigenvector for the system which produces the maximum transmitted irradiance has an eigenvalue equal to unity. It can be shown that this condition results in the relation

$$M_{11} + (M_{12}^2 + M_{13}^2 + M_{14}^2)^{1/2} = 1,$$

being satisfied, and that the overall loss by the system is accounted for by a scalar loss factor multiplying each matrix. This factor has not been determined in the present work.

Note also, that these matrices, like the Jones matrices from which they were derived, do not include depolarizing effects, as explained earlier.

B. System Calibration.

The problem of calibrating the telescope-spectroheliograph combination for use as an accurate polarimeter is made difficult by a number of factors. Ideally, one would like to have sources of known polarization located at a great distance from the telescope and make a complete end to end calibration by comparing the measured Stokes vector S' with the known incident vector S . Repeating this for four independent incident vectors one could then solve for the elements of the overall system matrix. Since not all 16 matrix elements are independent this procedure would overdetermine the matrix, but the redundancy can be used as a check on the accuracy achieved.

While it may be possible to carry out such a procedure for a small number of telescope pointings corresponding to local sites where a calibrated source could be set

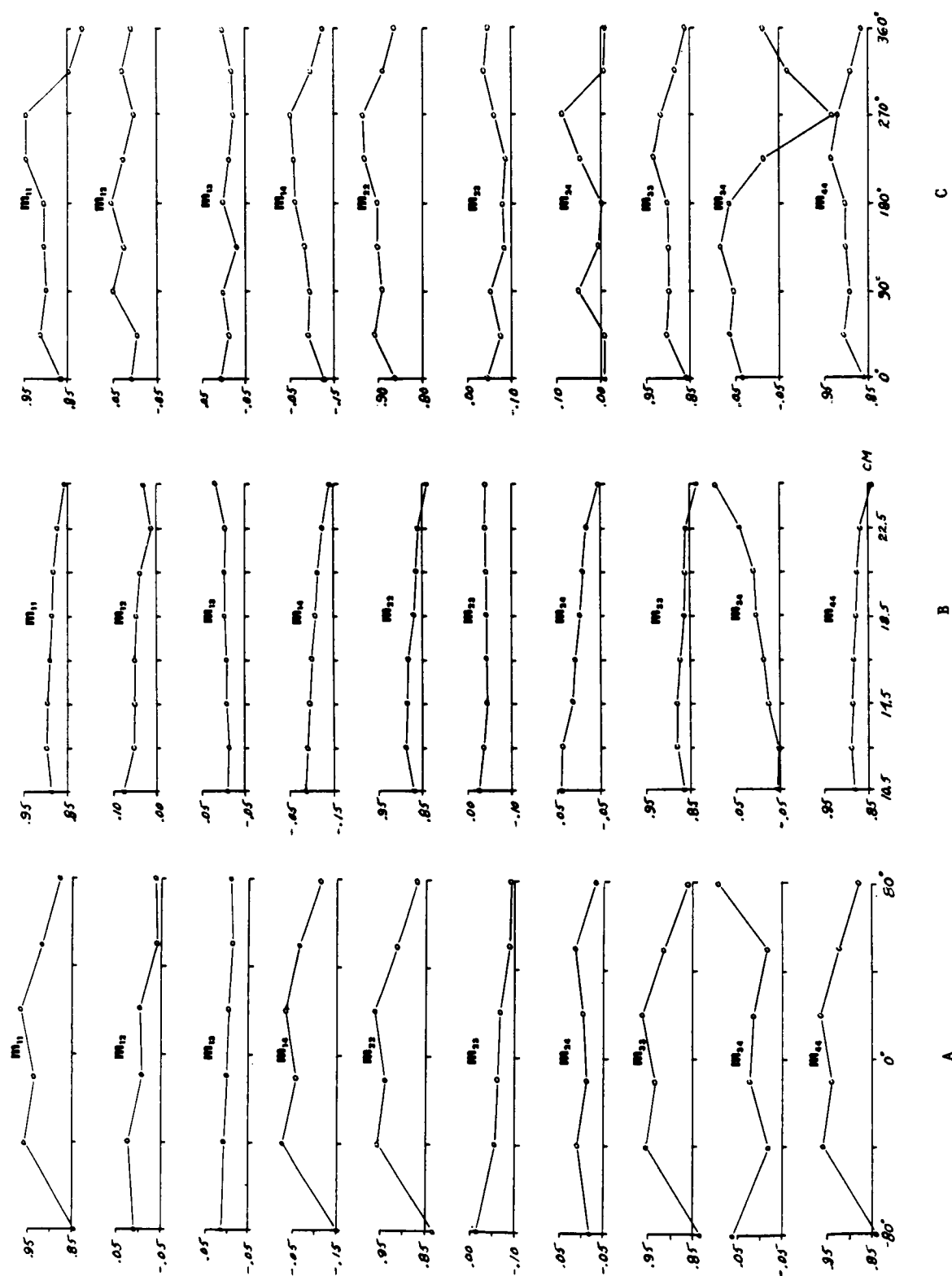


Figure 7: Telescope matrix elements as a function of hour angle (a), and position of the beam in the entrance window (b,c). The linear coordinate measures distance from a vertical reference axis located 8.4 cm below a horizontal reference axis on the window, while the angular coordinate represents measurements made at a constant radial distance of 20 cm from the center of the window.

up, say on a mountain top, this would not suffice to determine the system matrix for other values of hour angle and declination, which in our case is a necessity. Another alternative is the use of large area sheet polarizing materials directly in front of the entrance window, but it is doubtful that the homogeneity of such materials is adequate to achieve the required accuracy.

At the present time we are considering a method combining the mountaintop source concept discussed above with the use of small aperture beams in the entrance window. This approach, details of which have yet to be worked out, would be based upon computing a large aperture matrix from a prescribed array of small aperture measurements carried out across the face of the entrance window, using the theory of partial coherence and partial polarization. The results of such computations would then be checked against full aperture measurements using a tunable laser with precision polarizing optics mounted on a nearby mountain. Once theory and full aperture measurements are brought into agreement, the former can be used to determine the required system matrices for all telescope angles.

6. Conclusions

Measurements of the polarizing characteristics of the observatory's 61 cm aperture vacuum solar telescope and 3 m focal length vacuum spectroheliograph show that these are relatively small, indicating that compensation can be achieved by matrix inversion methods, thus permitting the instruments to be used as the image and spectrum forming parts of a Stokes Polarimeter for vector solar magnetic field and other solar studies.

A design study for such an instrument has been carried out based on the use of a TV camera system to scan the multiple images of the spectroheliograph exit slit formed by a combination of non-polarizing and polarizing beamsplitters. This arrangement produces six beams whose states of polarization are such that sums and differences of these result, after the application of suitable photometric and polarimetric corrections, in the four Stokes vector element profiles $I(\lambda)$, $Q(\lambda)$, $U(\lambda)$, $V(\lambda)$, each profile corresponding to a given solar image point. Vertical scanning of the TV camera samples a line of such image points corresponding to the position of the spectroheliograph entrance slit on the solar image, and mechanical scanning of this slit across the image permits the recording of Stokes profiles over a prescribed area of the solar disc.

The proposed design possesses the advantages of photometric efficiency, in that no light that passes through the spectroheliograph exit slit is sacrificed, aside from unavoidable small losses due to reflection, scattering and absorption, simplicity, in that no rotating devices or electro-optic modulators are used, and speed, in that the recording of the raw video signal is done in real time, i.e., at the standard TV rate of 30 frames/sec., using a standard VCR.

Off-line processing and data reduction of the analog video signal stored on the video cassette is accomplished by means of a real time image digitizing system which converts each video frame into an 8-bit, 256 x 256 pixel array, and permits the integration of up to 16 such frames in approximately one half second, and a computer, which reads the digitized arrays into memory line by line, and operates on these data with photometric and polarimetric calibration data stored on disc.

The final step in the data reduction process involves the calculation of weighted integrals of the Stokes profiles, from which vector magnetic field data can be extracted by comparison with theory. This integration process, in addition to

providing a 14 fold reduction in the data, improves the overall signal to noise ratio and greatly reduces the sensitivity of the measurements to spectrum line positioning errors and Doppler shifts.

The polarimeter design is also amenable to a number of obvious improvements that can be considered once the system is working, as for example, the replacement of the TV camera system with a two-dimensional diode array, and the extension of the measurements to more than one spectral line at a time. Use of an echelle grating in the spectroheliograph with a free spectral range of $\sim 120 \text{ \AA}$ permits a rapid change of wavelength by incorporating a rotating wedge prefilter into the design, thus permitting the production of sequential maps obtained with different lines.

Other improvements would include incorporating the 28 cm (11 inch) telescope image to permit increased area coverage at the cost of reduced resolution. This telescope system shares the same optical train as the 61 cm instrument except for the main image forming optics, and provides a reduced image size at the same focal ratio of $f/20$.

Finally, the polarimeter optics are simple and can readily be modified to include cylindrical components in order to provide an anamorphic system capable of permitting a choice of exit slit widths to be used.

All of the above is purchased at the cost of placing great demands on the TV system, with respect to stability of gain and image geometry, as well as on the calibration procedures required to compensate for photometric and polarimetric errors. Present indications are that these problems are soluble, and that the resultant system will provide high quality polarimetric data for a variety of important solar studies.

References

- Bashara, N. M. (ed.): 1969, *Surf. Sci.*, 16.
Gerrard, A. and Burch, J. M.: 1975, *Introduction to Matrix Methods in Optics* (John Wiley and Sons, N. Y.), 328.
Holmes, D. A.: 1964, *J. Opt. Soc. Am.*, 54, 1115.
Smith, P. H.: 1969, *Surf. Sc.*, 16, 34.
Zeldin, L. K.: 1981, Thesis, Dept. of Physics and Astronomy, Cal. State Univ., Northridge.

The new solar magnetograph for
the Canary Islands Observatory

T.H.E.M.I.S. *

BY

J. Rayrole

Observatoire de Meudon

THEMIS has been designed for accurate polarization measurements in order to determine the intensity and direction of the magnetic field without interference with the local variations of the thermodynamical parameters (temperature, pressure, density...).

That goal leads to the following requirements :

- 1) High spatial resolution in horizontal direction.
- 2) Sufficient resolution in height, which requires observations in several spectral lines with a sufficient spectral resolution.
- 3) Accurate polarization measurements.
- 4) Adequate time coverage to follow the evolution of individual structures.
- 5) Precise tracking and scanning mechanisms.
- 6) Sufficient field of view for the different structures we have to observe.

For the first condition, it is superfluous to tell solar physicists that high spatial resolution is the key of the modern solar physics.

It is the same for the second condition, but it is perhaps less evident for everybody that simultaneous observations with several spectral lines are necessary to be able to separate the effects of thermodynamical parameters (T , p , ρ) on magnetic or velocity field measurements.

*) Télescope Héliographique pour l'Etude du Magnétisme et des Instabilités Solaires.

The high gradient of magnetic field in the deepest part of the sun's atmosphere is correlated with high gradient of the thermodynamical parameters.

The size of the structure is so small that, as good as the resolution is, the atmospherical smoothing of the image cannot be entirely neglected.

So, to disconnect the magnetic field measurements from temperature and density variations, we must observe several spectral lines more or less sensitive to these parameters and that must be done at different levels in the sun's atmosphere.

Most of the structures are magnetic. Magnetic field is one of the most important parameters we have to observe but also one of the most difficult to obtain.

Magnetographs must be designed not only for accurate polarization measurements but also for correct determination of the intensity and direction of the magnetic field from observed parameters (calibration).

All these conditions cannot be obtained simultaneously.

For groundbased instruments, some desired specifications exclude some others. So we must make choices and optimize the instrument for its well defined goal.

For a limited diffraction telescope, the resolution increases with the aperture, but in fact, for a groundbased instrument, the resolution will be limited by the atmosphere.

Upper atmosphere seeing.

The influence on image quality of atmospheric turbulence can be represented by a single parameter which usually is referred to as Fried's parameter R_0 :

$$R_0 = 0.18 L^{6/5} (\cos(Z))^{3/5} \left(\int C_n^2(h) dh \right)^{-3/5}$$

where L is the wavelength and Z the angular zenith distance.

The structure constant of the refractive index, $C_n^2(h)$ describes the amount of refractive inhomogeneities as function of altitude. The real resolution will be the resolution of an ideal telescope of aperture R_0 .

The wavelength dependence of the index of refraction leads to variation of the coherence area with wavelength.

For Speckle interferometry (Fried) :

$$DL = 1.426 \lambda (R_0/D)$$

For a Michelson interferometer (Roddier) :

$$DL = 0.45 \lambda (R_0/D)^{5/6}$$

So the aperture of the telescope, for a given R_0 value, must be limited in function of the needed spectral range (Fig. 1).

For instrumental polarization point of view, from the entrance aperture to the end of the analyser of polarization, the optical scheme must have a symmetry of revolution. For any telescope of this kind, the occultation of the secondary mirror is about 15% and it will not be possible to limit the aperture by a mask at half its value without occulting all the beam.

So the aperture of the telescope must be adapted to the best R_0 value we should hope on the site, and the needed spectral range.

The choices for THEMIS.

- 1) Resolution of 0.1 arc sec for Speckle interferometry leads to $D > 90$ cm
- 2) Coherent spectral range $> 1.000 \text{ \AA}$ for :
 - $R_0 = 45$ cm
 - $L > 6.000 \text{ \AA}$
 - Zenith angular distance $< 75^\circ$
 - leads to $D < 90$ cm
- 3) Instrumental polarization as low as possible leads to D as small as possible?
- 4) Analyser of polarization designed in order to obtain the whole profiles for the different lines, simultaneous data reduction process different techniques.

Instrumental polarization

For each optical surface :

$$\begin{vmatrix} I \\ Q \\ U \\ V \end{vmatrix} = \begin{vmatrix} (T_x^2 + T_y^2)/2 & (T_x^2 - T_y^2)/2 & 0 & 0 \\ (T_x^2 - T_y^2)/2 & (T_x^2 + T_y^2)/2 & 0 & 0 \\ 0 & 0 & T_x T_y \cos B & T_x T_y \sin B \\ 0 & 0 & -T_x T_y \sin B & T_x T_y \cos B \end{vmatrix} * \begin{vmatrix} I_0 \\ Q_0 \\ U_0 \\ V_0 \end{vmatrix}$$

where I_0, Q_0, U_0, V_0 (incident Stokes parameters) are referred to the optical axes of the surface.

T_x = Coefficient of transmission or reflection in the incident plane

T_y = Coefficient of transmission or reflection in the perpendicular plane

B = Phase retardation

The optical surface will not perturb the state of polarization of the incident beam if :

$$\begin{aligned}
 T_x^2 - T_y^2 &< \epsilon_1 & \text{and } B &< \epsilon_2 \\
 \sin B &< \cos B/1000 & B &< 0.06^\circ
 \end{aligned}$$

For a mirror $T_x^2 - T_y^2$ and the B value equal 0 for normal incidence and increase with the incidence angle; the variations are functions of the coating.

For a vacuum telescope, the entrance and exit windows are subject to :

1) Mechanical stresses, because of the difference in pressure between its two sides and the reactions of its supports and the airtight joint.

2) Thermal stresses, because of the local difference of temperature due to radiation (sun, sky, structure...), conduction and convection (ambient air).

These local stresses, according to Brewster's law, introduce variations of glass index and hence optical phase retardations which are proportional to them. Only the mechanical reaction of the joint and the radial gradient of temperature introduce polarization.

The main results are the following :

1) the local mechanical or thermal polarization is equal to zero in the center and maximal on the rim.

2) the global mechanical or thermal polarization is equal to zero for a revolution symmetry distribution of the local stresses.

Entrance window

The main effect will be the global polarization. It is necessary to control the revolution symmetry of the reactions of its supports and the air-tight joint and the symmetry of radial temperature gradient.

Exit window

The main effect will be the local polarization. Its diameter must be large enough to have in the center a sufficient area with weak local polarization.

Principle of T.H.E.M.I.S.

A Ritchey-Chretien telescope pointed to the sun (Fig. 2) focuses a solar image on the spectrograph slit.

Behind it, we put a birefringent plate. The axes of this plate are parallel and perpendicular to the slit (Fig. 2 and 3). Then we put two calcite crystals, the sides of which are cut at 45° to the crystallography axis.

In relation one to the other, they are crossed and directed so that the two crossed linear beams they transmit have the same optical length and are polarized at 45° to the slit.

At this stage, the polarimetric analysis of the light is completed, so that the polarization state has the least interference with the instrument. Then nothing can change the shape of the line profiles if one can separate completely the two beams given by the analyzer at the spectrograph exit.

The predisperser characteristics and the echelle spectrograph allow this separation without any pollution.

The dispersion of the two spectrographs being correctly chosen by changing the distance between the two mirrors M_1 and M_2 , we are able to do that for any lines and simultaneously for a great number of lines (Fig. 4).

Structure of the instrument

For the framework, we adopt an entirely metallic structure both for the instrument and the tower. The vertical spectrographs are stiffly tied to an azimuthal mounting which bears the telescope (Fig. 5,6,)

The optical pieces will be cooled by heat pipes. A heat pipe is :

- A passive system for transporting large quantities of thermal energy
- It can have an effective thermal resistance several thousand times less than the best metallic conductors.
- It has no moving parts, requires no external power source and has an indefinite life.

A heat pipe consists of a closed container which has been internally lined with a "wick". The container is evacuated and the working fluid is leaked in, just enough fluid to saturate the wick.

If we have a difference of temperature, the vapour moves from the hot part to the cold part of the container and the fluid capillary returns via the wick.

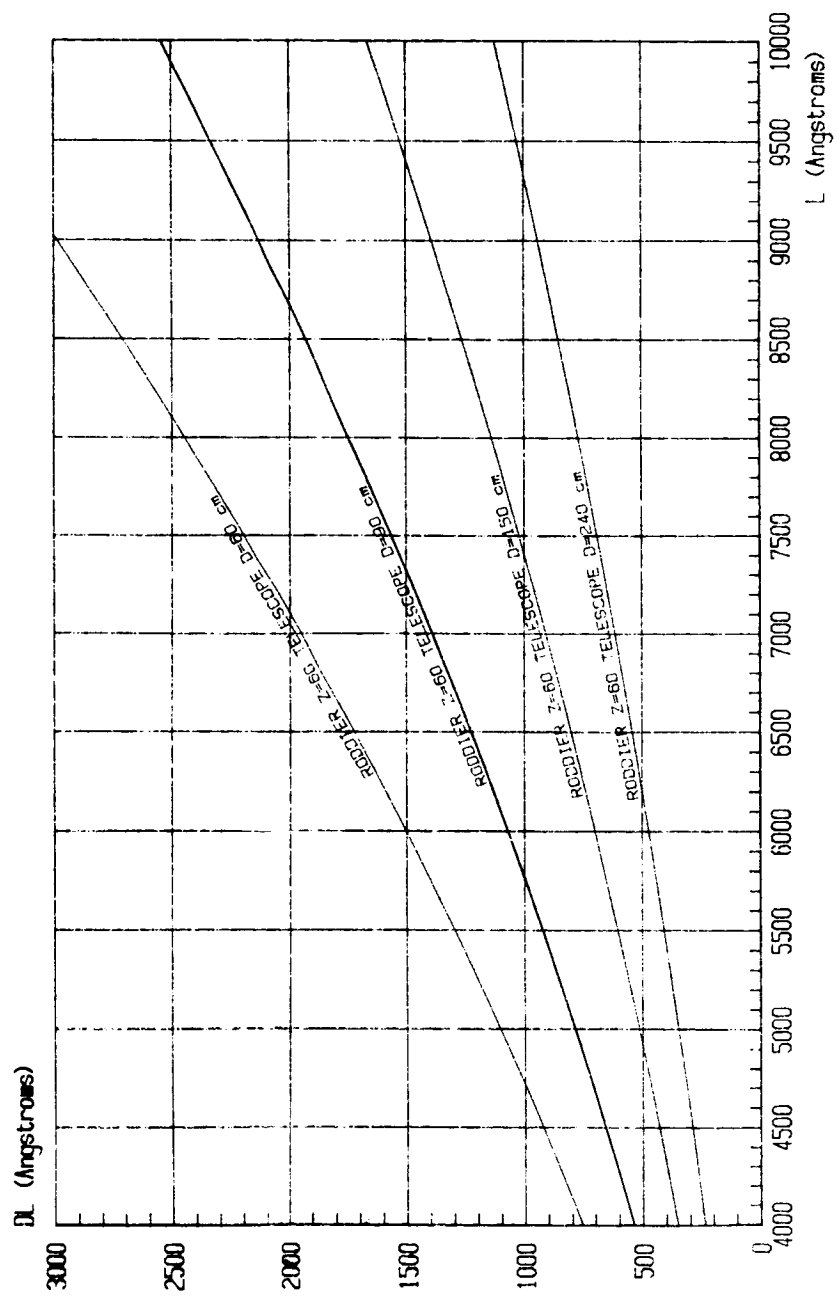
If we control electronically the difference of temperature, we can adjust the conductance of the heat pipe.

High resolution scanning of the sun's surface

For many astrophysical problems, we need good maps of the space or time derivatives of the observed parameters. The inertia of a telescope introduces a loss of efficiency of the drive mechanisms to scan the solar image rapidly with high accuracy.

Optical scheme of THEMIS has been designed so that such effects would be as small as possible. High accuracy scanning is obtained by moving the spectrograph slit in the focal plane of the telescope (F1, Fig. 5). A servo controlled flat mirror (image of the pupil) holds the beams on the predisperser optical axis.

With such a device, we are able to scan an area of 4' x 4' inside the 6' field of view.



$R_0(6000) = 45$ cm for $Z=0$

Figure 1

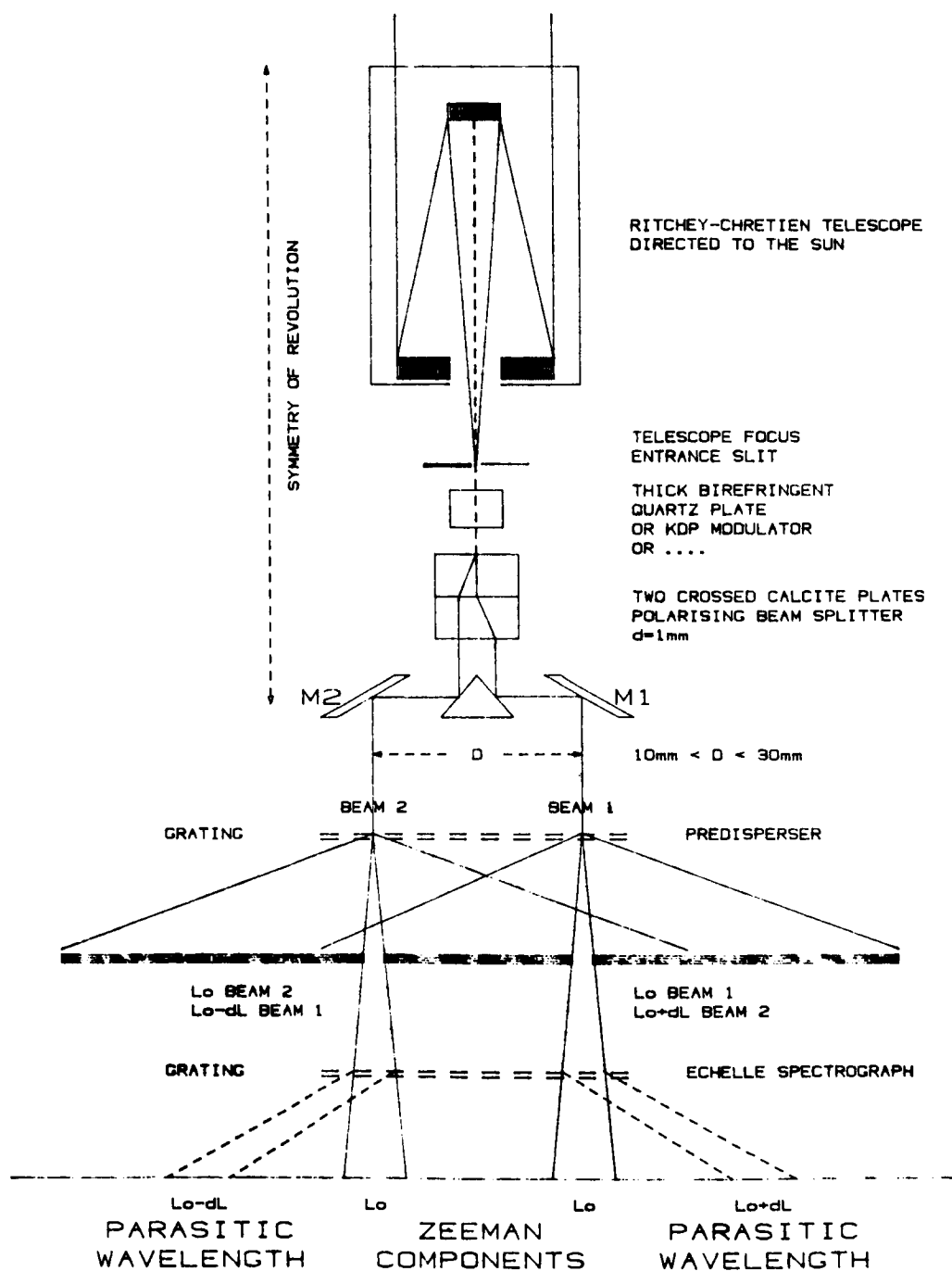


Figure 2

T.H.E.M.I.S. ANALYSER OF POLARIZATION

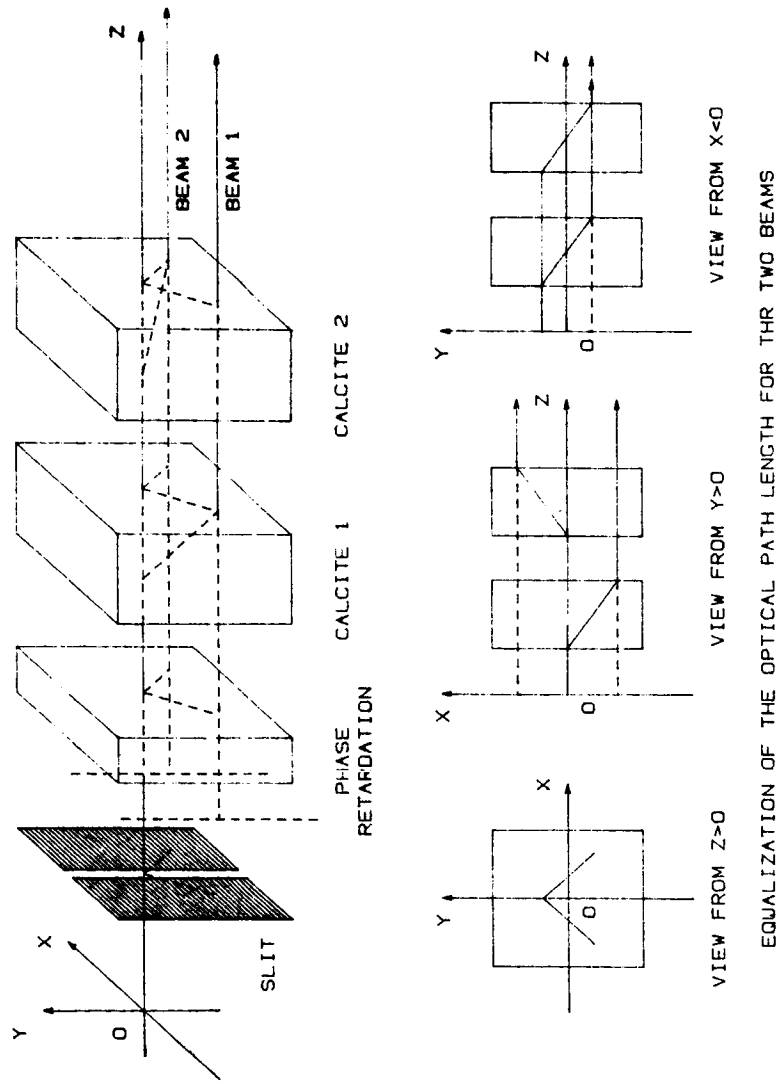


Figure 3

Figure 4 Generally most of the parasitic wavelength fall in the continuum. They are useful for simultaneous observations in continuum and spectral lines. They are also useful to control the internal coherence of the used spectral range.

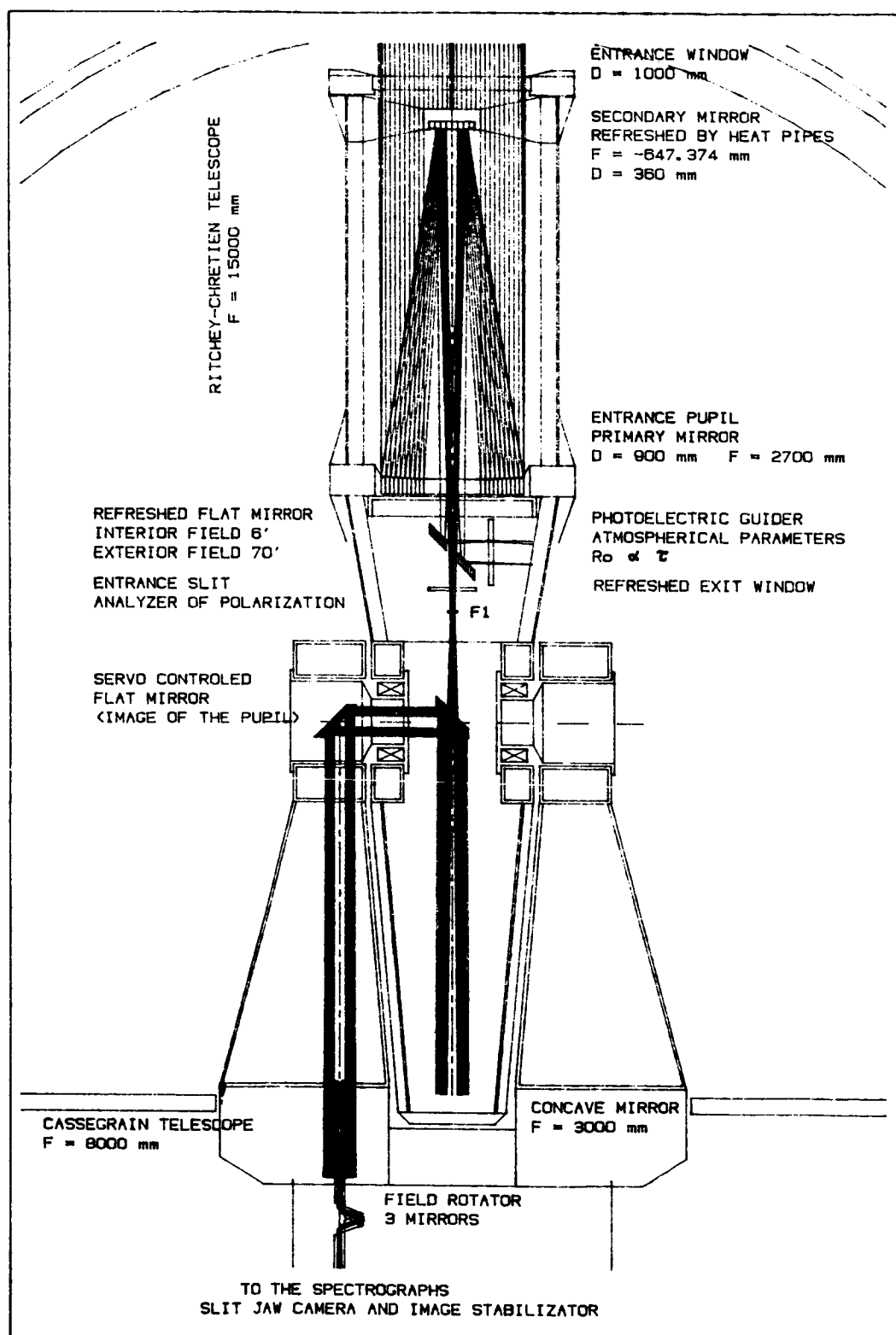
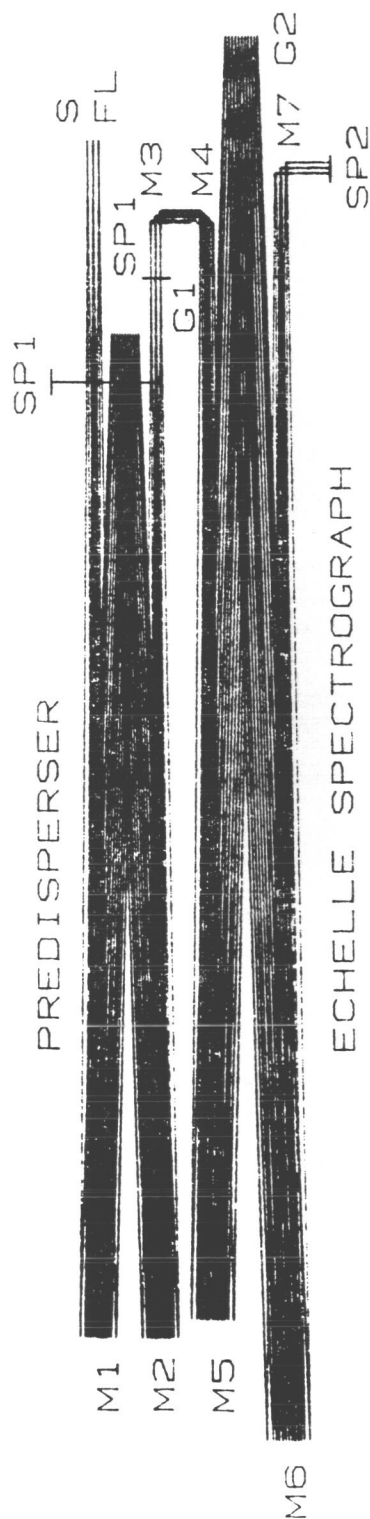


Figure 5



PREDISPERSER :

S ENTRANCE SLIT
 FL FIELD LENS
 M1 COLLIMATOR MIRROR
 $F = 7900 \text{ mm}$
 G1 PREDISPERSER GRATING
 150 g/mm 1st ORDER.
 BLAZE 2 DEGRES
 1200 g/mm 2nd ORDER
 BLAZE 42 DEGRES
 79 g/mm BLAZE 63 DEGRES
 M2 CAMERA MIRROR
 $F = 6300 \text{ mm}$

SP1 PREDISPERSER FOCUS
 0.084 mm/A
 2 mm/A
 5 mm/A
 SP1' ALTERNATE EXIT

ECHELLE SPECTROGRAPH

M3 FLAT MIRROR
 M4 FLAT MIRROR
 M5 COLLIMATOR MIRROR
 $F = 7500 \text{ mm}$
 G2 ECHELLE GRATING
 79 g/mm BLAZE 63 DEGRES
 M6 CAMERA MIRROR
 $F = 8000 \text{ mm}$

SP2 EXIT FOCUS
 4 TO 5 mm/A

Figure 6

The Vector Magnetograph of the Sayan Solar Observatory

V.M.Grigoryev, N.I.Kobanov, B.F.Osak, V.L.Selivanov, V.E.Stepanov/Sayan Solar Observatory, Siberian Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation, Siberian Dept., Academy of Sciences, Irkutsk, U.S.S.R.

1. Introduction

By the mid-70's at the Sayan-Mountain Solar Observatory we devised and constructed a new solar magnetograph, designed to provide measurements of the vector magnetic field of the Sun. In the developmental stage we tried to take into consideration our observational practice with the earlier magnetograph of the Sayan Observatory (Kuznetsov et al., 1966) as well as that of other researchers, primarily Livingston and Harvey (1971).

One of the major challenges is the problem of choosing an optimum scheme for the electrooptical analyzer of polarization and related control principles. These questions are examined in Section 3. Section 4 describes an electrooptical deflector, which is advantageously used to employ a single photodetector and to remove systematic errors inherent in magnetographs with two photodetectors in the wings of the line. Adjustment errors of optical elements of the polarization analyzer and errors of control voltages are discussed in Section 5. A method for measuring the telescope's polarization matrix, reduction of magnetographic measurements for instrumental polarization and calibration of magnetograph channels are discussed in Section 6. Section 7 is devoted to questions of evaluating scattered light and of reducing magnetic field measurements. A computational procedure for the magnetic field vector parameters is briefly outlined in Section 8. Section 9 presents an outline of the computer control of the solar magnetograph and of the processing and control system software.

2. The Instrument Layout

The vector magnetograph is installed at the horizontal solar telescope of the Sayan Solar Observatory (Osak et al., 1979). The general scheme of the telescope and magnetograph is shown in Fig. 1. A two-mirror coelostat system feeds the primary mirror of the telescope, 800 mm in diameter and of a 20 m focal length. The center of the primary mir-

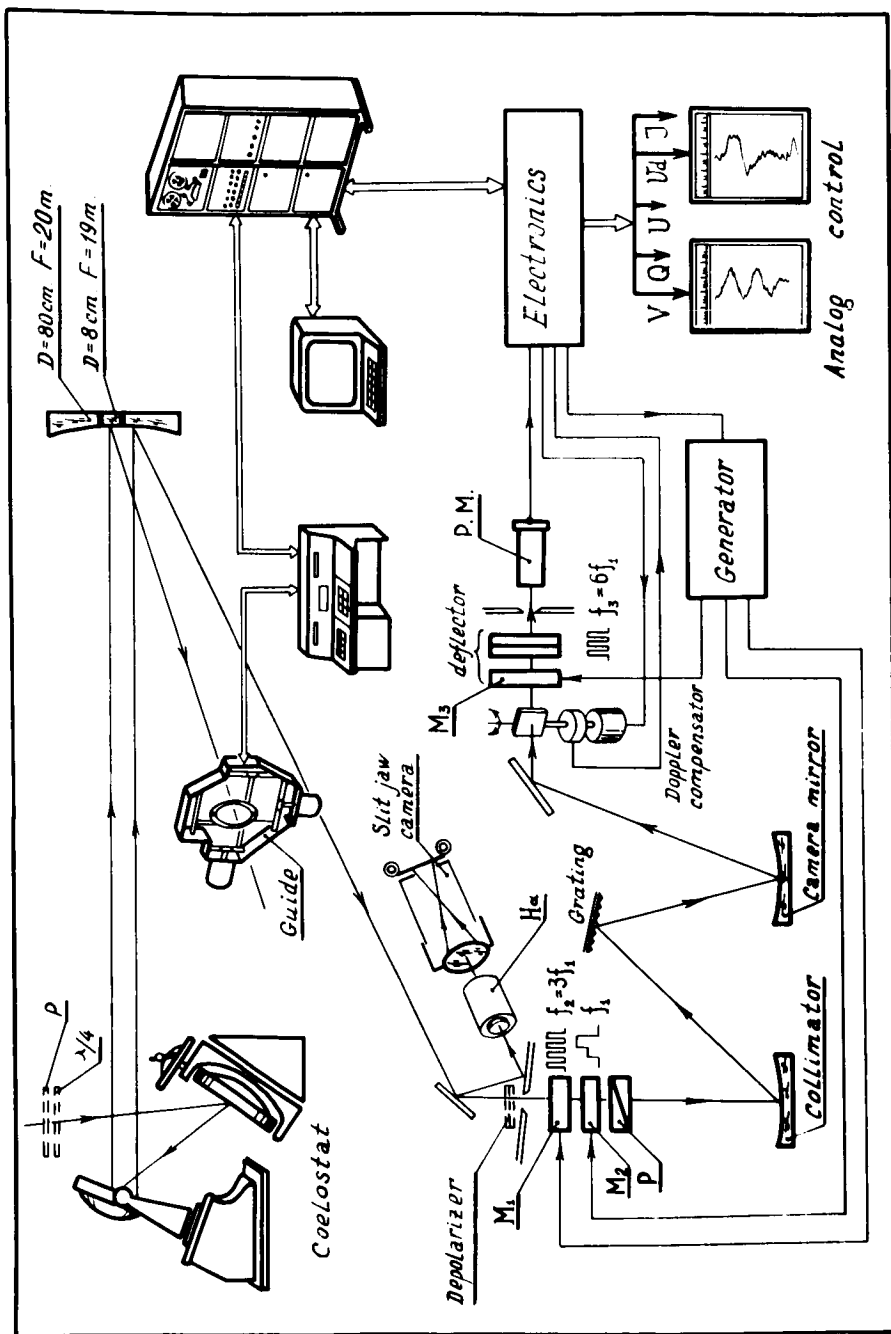


Fig. 1. The general scheme of the telescope and magnetograph of the Sayan Solar Observatory

ror is provided with an opening 100 mm in diameter, in which is placed the spherical mirror of the tracking system, 80 mm in diameter and of 10 m focal length. The two mirrors are made of glass ceramic and the small mirror is secured inside of the primary with glass ceramic wedges. The choice of material with a low coefficient of thermal expansion and the fastening system of the small mirror preclude a possible misalignment of the two optical systems. The inclination of the small mirror to the primary is about 1° . By means of a flat diagonal mirror the primary mirror forms a solar image on the spectrograph entrance slit. The light beam, reflected from the smooth slit jaws, is used to generate H-alpha or K Ca II filtergrams by means of birefringent filters and a photographic camera.

The small mirror forms a solar image in the plane of the tracking system detectors. These are installed on a movable carriage that can be displaced along two orthogonal directions by means of micrometer screws and stepping motors. The tracking system correction drives are installed on the second flat mirror of the coelostat system. These drives employ stepping motors and a reduction gear of "wedge-lever" type. The principal characteristics of the tracking and scanning system are: a 1 arc sec guiding accuracy, scanning steps of 0.5 arc sec each, and the image scanning rate ranges from 4 arc sec per hour to 30 arc sec per sec.

The spectrograph is a horizontal one and the collimator and camera mirror are of a 7 m focal length. The diffraction grating, whose dimensions are 200 x 200 mm, has 600 grooves per 1 mm and the blaze angle corresponds to the fifth order in the green. The dispersion in the exit slit plane is about 3 mm/A in the fifth order in the green.

The spectrograph entrance slit is followed by an electrooptical polarization analyzer (EOPA) that includes two DKDP crystals and a linear polarizer. In the spectrograph focal plane is placed a photometer. The photometer design employs a single photomultiplier, unlike conventional systems with two photomultipliers. In order to provide measurements in the two wings of the absorption line, the exit slit is preceded by an electrooptical deflector (EOD). This consists of a DKDP crystal and a calcite plate. Due to the effect of a square-wave voltage fed to the DKDP crystal, the deflector alternately switches the photometer from one wing of the line to the other.

Line-of-sight velocity measurements and Doppler shift compensation are effected by rotating the plane-parallel plate. The plate is rotated by means

of a stepping motor and the position of the plate is read by a shaft encoder. The error in the velocity measurements is ≈ 30 m/sec and the integration time is ≈ 0.3 sec.

3. The Electrooptical Polarization Analyzer

In the general case an electrooptical polarization analyzer (EOPA) can be represented by a combination of two variable phase plates (modulators M_1 and M_2) and a polarizer P (Fig. 2). The effect of the analyzer on polarized light can be conveniently described in terms of the Muller matrix (Shurcliff, 1965). Let the coordinate system be coincident with the crystal M_2 axes. The axes of crystal M_1 and polarizer P are rotated with respect to those of crystal M_2 through angles β and α , respectively. The effect of the entire system on polarized light is described by the Muller matrix:

$$M = P(\alpha)M_2(\delta_2, 0)M_1(\delta_1, \beta). \quad (1)$$

If a radiation with the parameters $\{I, Q, U, V\}$ is incident on the system, then it may be demonstrated that the intensity of the resulting radiation is defined by the following expression:

$$\begin{aligned} I' = & \frac{1}{2}I + \frac{1}{2}\left\{[1 - (1 - \cos\delta_1)\sin^2 2\beta]\cos 2\alpha + \left[\frac{1}{2}(1 - \cos\delta_1)\sin 4\beta \cos \delta_2 + \right. \right. \\ & \left. \left. + \sin\delta_1 \sin\delta_2 \sin 2\beta\right]\sin 2\alpha\right\}Q + \frac{1}{2}\left\{\frac{1}{2}(1 - \cos\delta_1)\sin 4\beta \cos 2\alpha + \right. \\ & \left. + [\cos\delta_1 \cos\delta_2 + (1 - \cos\delta_1)\cos\delta_2 \sin^2 2\beta - \sin\delta_1 \sin\delta_2 \cos 2\beta]\sin 2\alpha\right\}U + \\ & + \left\{\frac{1}{2}\sin\delta_1 \sin 2\beta \cos 2\alpha - [\sin\delta_1 \cos\delta_2 \cos 2\beta + \cos\delta_1 \sin\delta_2]\sin 2\alpha\right\}V, \end{aligned} \quad (2)$$

where δ_1 and δ_2 are the phase shifts, introduced by crystals M_1 and M_2 , respectively. On substitution of specific values of the angles α and β and of phase shifts $\delta_1(t)$, $\delta_2(t)$ into the time functions, it is possible to consider the various systems of electrooptical polarization analyzers which were introduced in solar magnetographs (Kotlyar, 1961; Stepanov and Severny, 1962; Ioshpa and Obridko, 1964; Kuznetsov et al., 1966; Livingston and Harvey, 1971). A common feature of the above-mentioned magnetographs was the use of sinusoidal control voltage on electrooptical crystals ($\delta = \delta_0 \sin \omega t$) where the information about the parame-

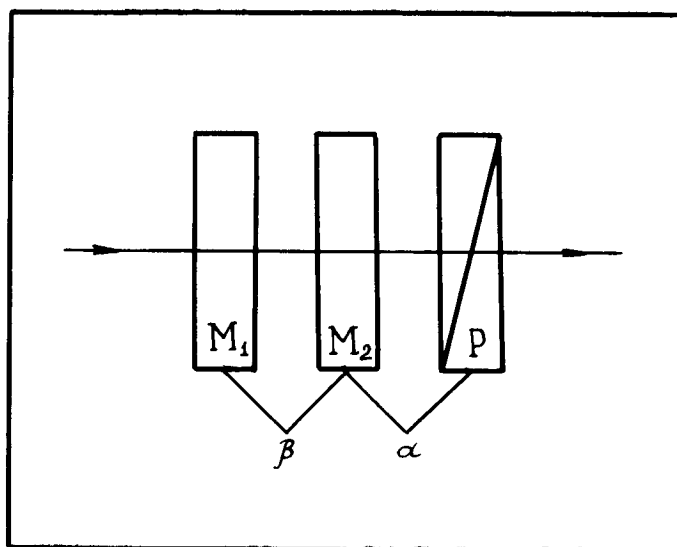


Fig. 2. An electrooptical polarization analyzer (General case)

ters Q , U and V is contained in the harmonics of the signal at the frequencies ω and 2ω . In order to separate the signals, the property of evenness of the parameters Q and U and of oddness of the parameter V with respect to the spectral line center is used. Then, addition and subtraction of signals from the two wings of the line make it possible to separate the three parameters. This procedure is applicable for the case of a symmetric line and a symmetric disposition of the slits relative to the line center. Otherwise, there occurs cross-talk between the parameter V and the parameter Q signal and vice versa. This will add errors to magnetic field vector measurements.

In addition, minor distortions in the shape of sinusoidal voltage on the electrooptical crystal lead to an appreciable redistribution of harmonic power and to a possible appearance of false signals (Kobanov, 1979). Since the voltage amplitude on electrooptical crystals reaches 3-4 kV, nonlinear distortions of about 5-7% are nearly unavoidable.

The afore-mentioned disadvantages led us to conclude that it is necessary to reject sinusoidal modulation and that an optimum solution should be sought for in using rectangularly pulsed voltages. There are also general theoretical arguments in favor of rectilinear modulation in optical instruments (Shestov, 1967).

The optical system of the analyzer we use to date is shown in Fig. 3. Crystal M_1 is oriented at an angle $\beta = 22^\circ.5$ with respect to the crystal M_2 axes and polarizer P is oriented at an angle $\alpha = 45^\circ$. Substitution of these values into (2) yields the following expression for the resulting intensity:

$$\begin{aligned}
 I' = & \frac{1}{2} \left\{ I + \left[\frac{\sqrt{2}}{2} \sin \delta_1 \sin \delta_2 + \frac{1}{2} (1 - \cos \delta_1) \cos \delta_2 \right] Q + \right. \\
 & + \left[\cos \delta_1 \cos \delta_2 + \frac{1}{2} (1 - \cos \delta_1) \cos \delta_2 - \frac{\sqrt{2}}{2} \sin \delta_1 \sin \delta_2 \right] U \\
 & \left. - \left[\frac{\sqrt{2}}{2} \sin \delta_1 \cos \delta_2 + \cos \delta_1 \sin \delta_2 \right] V \right\}.
 \end{aligned} \quad (3)$$

The electrooptical crystal M_2 is fed with a staircase voltage of frequency f_0 , that consecutively produces within the crystal phase shifts 0 , $\lambda/4$, and $\lambda/2$ while crystal M_1 is fed with rectangular voltage of frequency $f_1 = 3f_0$ that produces phase shifts 0 and $\lambda/2$. Fig. 3b presents a time history of phase shifts on modulators M_1 , M_2 and deflector D.

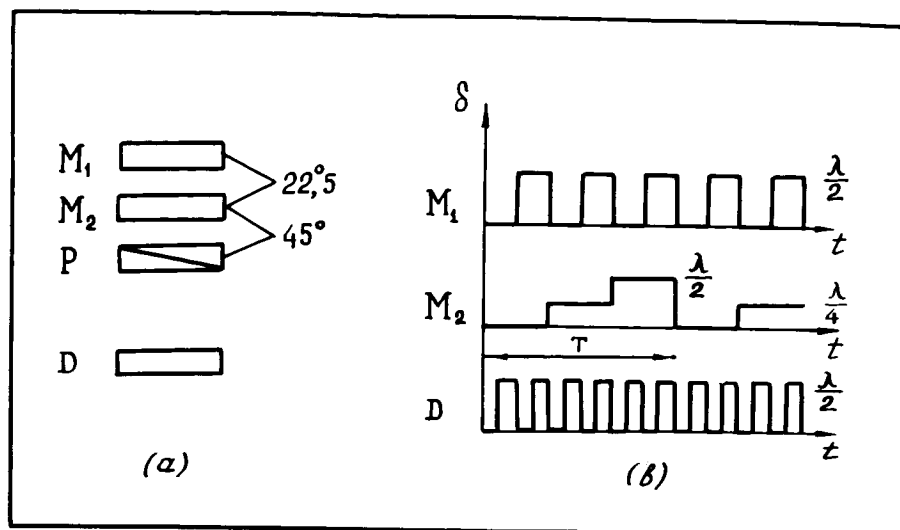


Fig. 3. a) The optical system of the polarization analyzer
b) The time history of phase shifts, produced by electrooptical crystals in such a system

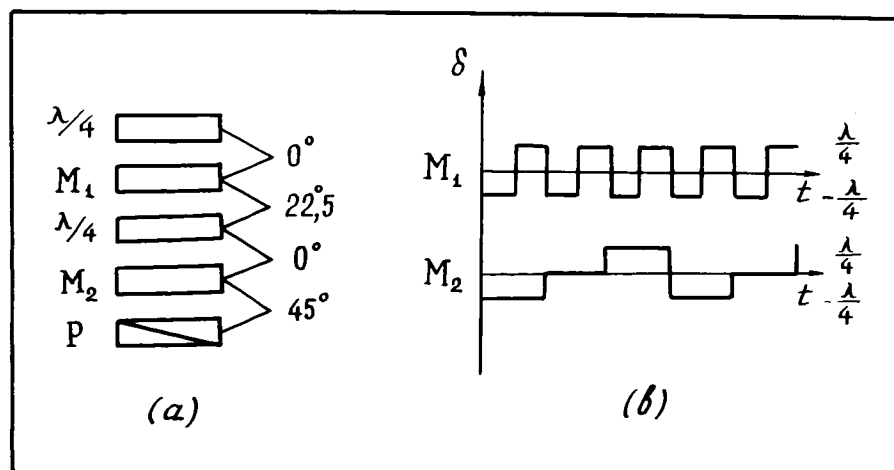


Fig. 4. a,b Same as in Fig. 3, ab but for an analyzer system version with composite elements

For a period $T = \frac{1}{f}$ it is possible to separate six different combinations of phase shifts:

1. $y'(0,0) = \frac{1}{2} (y+u),$
2. $y'(\frac{\lambda}{2},0) = \frac{1}{2} (y+Q),$
3. $y'(\frac{\lambda}{2},\frac{\lambda}{4}) = \frac{1}{2} (y-V),$
4. $y'(0,\frac{\lambda}{4}) = \frac{1}{2} (y+V),$
5. $y'(0,\frac{\lambda}{2}) = \frac{1}{2} (y-u),$
6. $y'(\frac{\lambda}{2},\frac{\lambda}{2}) = \frac{1}{2} (y-Q).$

(4)

Taking appropriate differences and integrating them over the exit slit width between the limits from λ_1 to λ_2 we obtain signals of Stokes parameters related to magnetic field parameters by the known relationships

$$\begin{aligned}
 S_Q &= \int_{\lambda_1}^{\lambda_2} [y'(\frac{\lambda}{2},0) - y'(\frac{\lambda}{2},\frac{\lambda}{2})] d\lambda = y_0 \varphi(H, \gamma) \cos 2\chi, \\
 S_u &= \int_{\lambda_1}^{\lambda_2} [y'(0,0) - y'(0,\frac{\lambda}{2})] d\lambda = y_0 \varphi(H, \gamma) \sin 2\chi, \\
 S_V &= \int_{\lambda_1}^{\lambda_2} [y'(0,\frac{\lambda}{4}) - y'(\frac{\lambda}{2},\frac{\lambda}{4})] d\lambda = y_0 f(H, \gamma),
 \end{aligned}
 \tag{5}$$

where y_0 is the continuum intensity and $\varphi(H, \gamma)$ and $f(H, \gamma)$ are the calibrating functions:

$$\begin{aligned}
 \varphi(H, \gamma) &= \frac{\int_{\lambda_1}^{\lambda_2} z_Q(H, \gamma) d\lambda}{\int_{\lambda_1}^{\lambda_2} z_y d\lambda}, \\
 f(H, \gamma) &= \frac{\int_{\lambda_1}^{\lambda_2} z_V(H, \gamma) d\lambda}{\int_{\lambda_1}^{\lambda_2} z_y d\lambda}.
 \end{aligned}
 \tag{6}$$

Contours of the polarization parameters z_y , z_Q , and z_V are calculated for given values of H , γ , and χ and for parameters of the spectral line and model atmosphere, $\varphi(\tau)$, $B(T(\tau))$, $\alpha(\tau)$ and $\Delta\lambda_D(\tau)$.

by numerical integration of a set of transport equations using the Runge-Kutta method. A detailed computational technique may be found in Katz, (1973) Grigoryev and Katz (1975).

In observational practice we are frequently faced with the need for a swift switch-over from magnetic field vector measurements to measurements of only the longitudinal component of the magnetic field and vice versa. Fig. 4a,b shows an analyzer scheme, in which for purposes of easing switch-over from the regime of field vector measurement to that of longitudinal component measurement as well as of reducing modulating voltages, phase elements M_1 and M_2 are each composed of a pair of an aligned $\lambda/4$ plate and an electrooptical crystal. Thus, crystal M_1 will be fed with rectangular voltage that produces a phase shift of $-\lambda/4$ and $+\lambda/4$, and M_2 is fed with staircase voltage that produces a phase shift of $-\lambda/4$, 0, and $+\lambda/4$. Switch-over to the regime of measurement of only the longitudinal component of the magnetic field vector is accomplished by switching off the modulating voltage on crystal M_2 .

4. The Electrooptical Deflector

Measurement of line-of-sight velocities and Doppler shift compensation require that the intensity be measured in the two wings of the line. Two phototubes are usually used in the wings of the line. The outputs of the two phototubes are balanced in the continuum before the start of the observation. A drift in the sensitivity in one phototube relative to the other will cause an error in line position. A 0.1 % balance error will induce a velocity error of tens of meters per second. Such errors are absent in the scheme with a single phototube, which uses an electrooptical deflector to switch over the photometer from one wing of the line to the other. A scheme of a photometer with an electrooptical deflector is shown in Fig. 5. The electrooptical deflector consists of an electrooptical crystal and two calcite plates, whose axes are mutually perpendicular and make angles of $\pm 45^\circ$ with the crystal axes. The deflector is placed before the photometer slit in the focal plane of the spectrograph. The radiation incident on the deflector past the electrooptical analyzer is linearly polarized. The deflector is fed with a modulating voltage of frequency δf , that produces a phase difference on crystal 0, $\lambda/2$. With phase 0 on the deflector the slit transmits light from one wing of the line while with phase $\lambda/2$, from the other wing. The separation of the portions of the line wings covering the photometer slit is

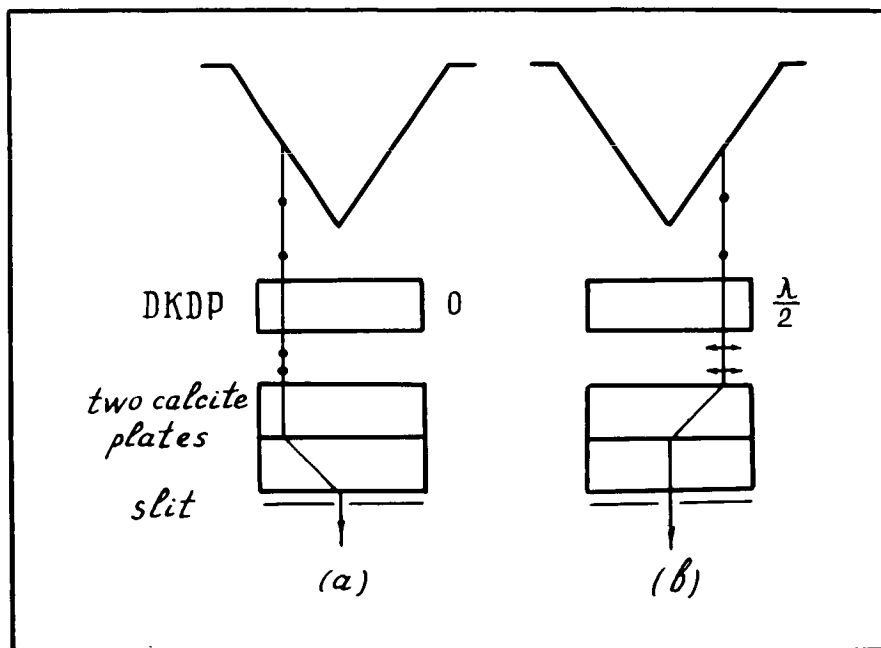


Fig. 5. Electrooptical deflector

determined by the thickness of calcite plates. The difference of signals measured in the right and left wing of the line for the period T_0 is proportional to the amount of Doppler shift of the line.

5. Errors of the Electrooptical Polarization Analyzer

5.1. Adjustment errors of optical elements of the analyzer and control voltages

Examine now a more general case of an analyzer with composite phase elements (Fig. 4a), assuming that $\lambda/4$ -plates are executed with high precision and only errors in the orientation of their axes are taken into account. The function of such an analyzer is described in terms of the matrix:

$$M = P(\alpha) M_2(\delta_2, 0) L_2(\frac{\lambda}{4}, x) M_1(\delta_1, \beta) L_1(\frac{\lambda}{4}, z), \quad (7)$$

where $L_1(\frac{\lambda}{4}, z)$, $L_2(\frac{\lambda}{4}, x)$ are the Muller matrices of the $\lambda/4$ -plates,

δ_1, δ_2 are the phase differences, introduced by crystals M_1 and M_2 ,
 x is the angle between the axes of a $\lambda/4$ -plate and crystal M_2 ,
 z is the angle between the axes of a $\lambda/4$ -plate and crystal M_1 ,
 β is the angle between the axes of crystals M_1 and M_2 , and
 α is the angle between polarizer P and the crystal M_2 axes.

The basic coordinate system coincides with the crystal M_2 axes. If the radiation with the parameters $\{I, Q, U, V\}$ is incident on the analyzer, then multiplying by matrix (7) yields an expression for the resulting intensity I' . We will not write it here because it is very complicated. On differentiating the expression for resulting intensity with respect to $\delta_1, \delta_2, \alpha, x, \beta$, and z and substituting the following values of angles: $\alpha = 45^\circ$, $x = 0^\circ$, and $\beta = z = 22^\circ.5$ we arrive at an expression for derivatives:

$$\begin{aligned} \frac{\partial I'}{\partial \delta_1} = & -\frac{1}{4} [\sqrt{2} \cos \delta_2 \sin \delta_1 + \sin \delta_2 \cos \delta_1] Q + \\ & + \frac{1}{4} [\sqrt{2} \sin \delta_1 \cos \delta_2 + \cos \delta_1 \sin \delta_2] U + \\ & + \frac{1}{2} [\sin \delta_1 \cos \delta_2 - \frac{\sqrt{2}}{2} \sin \delta_1 \sin \delta_2] V; \end{aligned} \quad (8)$$

$$\frac{\partial y'}{\partial \delta_2} = -\frac{1}{4}[\sqrt{2} \cos \delta_1 \sin \delta_2 + \cos \delta_2 (1 + \sin \delta_1)] Q + \frac{1}{4}[\sqrt{2} \cos \delta_1 \sin \delta_2 - \cos \delta_2 (1 - \sin \delta_1)] U + \\ + [\frac{\sqrt{2}}{2} \cos \delta_1 \cos \delta_2 - \cos \delta_1 \sin \delta_2] V;$$

$$\frac{\partial y'}{\partial \alpha} = -\frac{1}{2}(1 - \sin \delta_1) Q - \frac{1}{2}(1 + \sin \delta_1) U - \frac{\sqrt{2}}{2} \cos \delta_1 V;$$

$$\frac{\partial y'}{\partial x} = \frac{1}{2}[2 \cos \delta_2 + \sin \delta_2 (1 - \sin \delta_1)] Q + \frac{1}{2}[2 \cos \delta_2 + \sin \delta_2 (1 + \sin \delta_1)] U + \\ + \frac{\sqrt{2}}{2} \cos \delta_1 \sin \delta_2 V;$$

$$\frac{\partial y'}{\partial \beta} = \frac{1}{2}[\sqrt{2} \cos \delta_2 \sin \delta_1 - \sin \delta_2 (1 - \cos \delta_1 - \sin \delta_1)] Q + \\ + \frac{1}{2}[\sqrt{2} \cos \delta_2 \sin \delta_1 - \sin \delta_2 (1 - \cos \delta_1 + \sin \delta_1)] U + \frac{\sqrt{2}}{2} \sin \delta_2 (1 - \cos \delta_1) V;$$

$$\frac{\partial y'}{\partial z} = \frac{1}{2}[\sqrt{2} \cos \delta_2 (\cos \delta_1 - \sin \delta_1) - \sin \delta_2 (\cos \delta_1 + \sin \delta_1 - 1)] Q + \\ + \frac{1}{2}[\sqrt{2} \cos \delta_2 (\cos \delta_1 - \sin \delta_1) - \sin \delta_2 (1 + \cos \delta_1 + \sin \delta_1)] U - \frac{\sqrt{2}}{2} \sin \delta_2 V.$$

On substituting into expressions (8) the values of δ_1 and δ_2 , corresponding to the six combinations of phase shifts on crystals M_1 and M_2 , we derive representations for total errors in the determination of intensity for appropriate states of the polarization analyzer:

$$\begin{aligned} 1. \Delta y'(-\frac{1}{4}, -\frac{1}{4}) &= \frac{-\sqrt{2}}{4} V \Delta \delta_1 - \frac{1}{2} V \Delta \delta_2 - Q \Delta x - Q \Delta \alpha + (Q - \frac{\sqrt{2}}{2} V) \Delta \beta + (\frac{\sqrt{2}}{2} V - Q) \Delta z, \\ 2. \Delta y'(\frac{1}{4}, -\frac{1}{4}) &= \frac{\sqrt{2}}{4} V \Delta \delta_1 + \frac{1}{2} V \Delta \delta_2 - U \Delta x - U \Delta \alpha + (U - \frac{\sqrt{2}}{2} V) \Delta \beta + (U + \frac{\sqrt{2}}{2} V) \Delta z, \\ 3. \Delta y'(-\frac{1}{4}, 0) &= \frac{\sqrt{2}}{4} (Q - U) \Delta \delta_1 - \frac{1}{2} U \Delta \delta_2 + (Q + U) \Delta x - Q \Delta \alpha - \frac{\sqrt{2}}{2} (Q + U) \Delta \beta + \frac{\sqrt{2}}{2} (Q + U) \Delta z, \\ 4. \Delta y'(\frac{1}{4}, 0) &= \frac{-\sqrt{2}}{4} (Q - U) \Delta \delta_1 - \frac{1}{2} Q \Delta \delta_2 + (Q + U) \Delta x - U \Delta \alpha + \frac{\sqrt{2}}{2} (Q + U) \Delta \beta - \frac{\sqrt{2}}{2} (Q + U) \Delta z, \\ 5. \Delta y'(-\frac{1}{4}, \frac{1}{4}) &= \frac{\sqrt{2}}{4} V \Delta \delta_1 + \frac{1}{2} V \Delta \delta_2 + Q \Delta x - Q \Delta \alpha + (\frac{\sqrt{2}}{2} V - Q) \Delta \beta + (Q - \frac{\sqrt{2}}{2} V) \Delta z, \end{aligned} \quad (9)$$

$$6. \Delta \mathcal{Y}(\frac{1}{4}, \frac{1}{4}) = \frac{\sqrt{2}}{2} V \Delta \delta_1 - \frac{1}{2} V \Delta \delta_2 + U \Delta x - U \Delta \alpha + (\frac{\sqrt{2}}{2} V - U) \Delta \beta - (U + \frac{\sqrt{2}}{2} V) \Delta z.$$

Signals Q , U and V are formed by a pairwise subtraction of the relevant states of the analyzer and errors in their determination can be derived using (9):

$$\begin{aligned} \Delta V = & -(\frac{\sqrt{2}}{2} \Delta \delta_1 + \frac{1}{2} \Delta \delta_2 - \Delta \alpha - \sqrt{2} \Delta \beta + \sqrt{2} \Delta z) Q + \\ & + (\frac{\sqrt{2}}{2} \Delta \delta_1 + \frac{1}{2} \Delta \delta_2 - \Delta \alpha + \sqrt{2} \Delta \beta - \sqrt{2} \Delta z) U; \end{aligned} \quad (10)$$

$$\Delta Q = (\frac{\sqrt{2}}{2} \Delta \delta_1 + \Delta \delta_2 + \sqrt{2} \Delta z - \sqrt{2} \Delta \beta) V + 2(\Delta \beta + \Delta z - \Delta x) U;$$

$$\Delta U = -(\frac{\sqrt{2}}{2} \Delta \delta_1 + \Delta \delta_2 + \sqrt{2} \Delta \beta - \sqrt{2} \Delta z) V + 2(\Delta \beta + \Delta z - \Delta x) Q.$$

The error in determination of \mathcal{Y} will be the sum of errors (7) for period T:

$$\Delta \mathcal{Y} = -(\frac{1}{2} \Delta \delta_2 + 3 \Delta \alpha - 2 \Delta x) Q - (\frac{1}{2} \Delta \delta_2 + 3 \Delta \alpha - 2 \Delta x) U. \quad (11)$$

When measuring in the two wings of the line we perform an addition of the signals of the parameters Q and U and a subtraction of the parameter V signals. Taking into account the properties of evenness and oddness of Stokes parameters relative to the line center and using expressions (10) we will obtain the following errors in determination of S_Q , S_U and S_V :

$$\begin{aligned} \Delta S_Q &= 4(\Delta \beta + \Delta z - \Delta x) S_U, \\ \Delta S_U &= 4(\Delta \beta - \Delta z - \Delta x) S_Q, \\ \Delta S_V &= 0. \end{aligned} \quad (12)$$

These expressions are satisfied whenever the line is symmetrical and there is no Doppler shift. The Doppler shift signal D is equal to the intensity difference in the red and blue wings of the line $\mathcal{Y}^2 - \mathcal{Y}^0$ for period T, and the error in its determination is

$$\Delta D = 0. \quad (13)$$

For a polarization analyzer, consisting of two crystals M_1 and M_2 and of polarizer P (Fig. 3a), by adopting a similar procedure we can obtain the following expressions for errors:

$$\begin{aligned}\Delta S_Q &= 4\Delta\beta S_u, \\ \Delta S_u &= 0, \\ \Delta S_V &= 0, \\ \Delta D &= 0, \\ \Delta S_y &= -\left(\frac{1}{2}\Delta\delta_2 + 3\Delta\alpha\right)(S_Q + S_u).\end{aligned}\tag{14}$$

The obtained results lead to the following conclusions. In the case when Doppler line shifts are properly compensated for and the line profile is a symmetrical one, Stokes parameter measurements are not affected by errors in modulating voltages that determine the magnitude of phase shift in electrooptical crystals. Errors in the mutual orientation of the axes of the analyzer's optical elements have influence upon the measured results for the parameters Q and U only. Thus, an error in the angles α, β, γ , and δ of about $0^\circ.5$ induces a few percent relative error of measurement of the parameters Q and U . However, spectral lines generally show a noticeable asymmetry in the photosphere and, specifically, in a sunspot region. In this case measurements of Stokes parameters and, hence, of magnetic field parameters will be affected by errors in the magnitude of modulating voltages. Estimations and observational practice give evidence that the phase shift values, δ_1 and δ_2 on crystals should be maintained to within about 0.1% accuracy.

5.2. Control of the magnitude of phase shifts in electrooptical crystals

Electrically, an electrooptical crystal-based modulator is a capacitor, whose capacitance together with relevant cables is as high as 300 pF. A pulsed charge of the capacitor with a subsequent disconnection of same from the circuitry provide just the principle for our control voltage devices (Kobanov, 1974). These devices, executed based on pulse thyratrons, provide a quite short duration of voltage switching on crystals, of about $1 \cdot 10^{-6}$ sec with low power consumption.

It is known that the amount of voltage on a crystal required to provide a phase shift, e.g., of $\lambda/2$

depends on the modulated luminous flux wavelength, temperature, crystal's homogeneity, and on the modulator structure. It is for these reasons that the optical parameter itself, i.e., the amount of the required phase shift within the crystal, should be controlled, rather than voltage. We have offered a method for controlling a $\lambda/2$ -phase shift, which employs the symmetry property of the electrooptical shutter light characteristic (Grigoryev and Kobanov, 1977). The method permits easy control of modulating voltages prior to the observation and provides a better than 0.1% accuracy.

5.3. On the accuracy of alignment of the electro-optical crystal's z-axis relative to the spectrograph's optical axis

In optical systems of telescope and spectrographs, electrooptical polarization analyzers are usually placed in the paths of converging or diverging light beams with an up to 3° angular aperture. Hence, angular field effects become of major importance. The essential point here is that an electrooptical crystal does not represent a precisely $\lambda/4$ - or a $\lambda/2$ -plate for off-axis beams. In the polar coordinates, the curves of the same phase delay are described in terms of Cassini's ovals with poles at exit points of the crystal's optical axes. As the electric field within the crystal reverses direction, the plane in which optical axes lie rotates by 90° , with a concurrent rotation of the pattern of curves of the same phase delay. In this case, if the light beam cone axis is normal to the crystal, the phase delay integrated over all beams will be the same for the two phases of modulation. However, if the beam cone axis is inclined relative to the crystal, then the phase delay integrated over all beams will differ for different modulation states. This will induce a cross-talk of the signals of linear and circular polarization. Similar errors are introduced whenever there occurs a departure from circular shape of the light beam cone cross-section within the spectrograph. This may be attributable to a very oblique incidence of beams on the coelostat system mirror in the morning and evening hours or when the light beam cross-section is not inscribed in the diffraction grating but describes it. These effects were discussed by Grigoryev and Kobanov (1980), Duvall (1977), and Grigoryev and Ilgamov (1983), and the requirements were formulated as follows:

- 1) DKDP electrooptical crystals should have a thickness of 1 to 2 mm;

- 2) the optical system of telescope and spectrograph should provide a circularly shaped cross-section of the light beam cone, inscribed in the dimensions of the diffraction grating; and
- 3) the error in the alignment of the z-axis of an electrooptical crystal relative to the spectrograph's optical axis should not exceed 1/10 of the angular aperture of the light beam.

6. Reduction for the Instrumental Matrix of Telescope and Calibration Procedure

6.1. Instrumental polarization and line off-centering

The degree of instrumental polarization effect on magnetographic measurements is determined, on the one hand, by polarization properties of the entire instrument optics, and, on the other, by the type of polarization analyzer. The polarizing effect of the instrument is described in terms of a 4x4 Muller matrix $\mathbb{L}_T = \{\alpha_{ij}\}$. Then, the Stokes parameters of incoming radiation, $\mathbb{L}_0 = \{I_0, Q_0, U_0, V_0\}$, are transformed by the telescope into:

$$\mathbb{L} = \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \mathbb{L}_T \mathbb{L}_0 = \begin{pmatrix} \alpha_{11} I_0 + \alpha_{12} Q_0 + \alpha_{13} U_0 + \alpha_{14} V_0 \\ \alpha_{21} I_0 + \alpha_{22} Q_0 + \alpha_{23} U_0 + \alpha_{24} V_0 \\ \alpha_{31} I_0 + \alpha_{32} Q_0 + \alpha_{33} U_0 + \alpha_{34} V_0 \\ \alpha_{41} I_0 + \alpha_{42} Q_0 + \alpha_{43} U_0 + \alpha_{44} V_0 \end{pmatrix} \quad (15)$$

First of all, it should be emphasized that since the parameter V_0 is an odd function relative to the line center, when a line-of-sight velocity compensator is employed the instrumental polarization leads to off-centering the line on the photometer slits (Jager, 1972; Stepanov and Severny, 1962; Grigoryev and Selivanov, 1978; Grigoryev et al., 1980). The false signal of line-of-sight velocity, associated with line miscentering in our magnetograph, is defined by the expression:

$$D \sim y^2 - y^6 = \alpha_{14} V_0. \quad (16)$$

For magnetographs with a sinusoidal modulation such a signal is (Jager, 1972)

$$D \sim (\alpha_{14} + \alpha_{34} J_0) V_0 \quad (17)$$

where J_0 is the Bessel's function.

The values of the coefficients α_{14} and α_{34} range

between $-0.002 - 0.005$ and $0.07 - 0.10$. The coefficient α_{44} is more than one order of magnitude smaller than the coefficient α_{34} , and hence in a magnetograph system with rectangular control voltage, off-centering effects due to instrumental polarization are substantially smaller than those in magnetographs with sinusoidal control voltage. Therefore, to a good approximation we can adopt $\alpha_{44} = 0$ for our magnetograph. Besides, when using a rectangular form of polarization analyzer control, as shown in Figs. 3 and 4, at the formation of an intensity signal in the blue and red it is possible to exclude analyzer's states 3 and 4 (cf. expression 4) associated with the circular polarization parameters. Such a technique precludes the possibility of line off-centering, even if $\alpha_{44} \neq 0$.

The Stokes parameters of the light transmitted by the telescope in the blue and red are represented by the expression (15) and differ only in the signs which are opposite before the terms involving the parameter V_0 . The magnetograph signals, formed from the two portions of the line contour, can then be written thus:

$$\begin{aligned} S_y &\sim (\alpha_{11} I_0 + \alpha_{12} Q_0 + \alpha_{13} U_0), \\ S_Q &\sim (\alpha_{21} I_0 + \alpha_{22} Q_0 + \alpha_{23} U_0), \\ S_U &\sim (\alpha_{31} I_0 + \alpha_{32} Q_0 + \alpha_{33} U_0), \\ S_V &\sim \alpha_{44} V_0. \end{aligned} \quad (18)$$

Thus, the coefficients of the telescope polarization matrix α_{14} , α_{24} , α_{34} and α_{41} , α_{42} , α_{43} are not involved in the formation of magnetograph signals, and the transformation of incident radiation I_0 into magnetograph signals S can be written in a matrix form

$$S = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & 0 \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & 0 \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & 0 \\ 0 & 0 & 0 & \alpha_{44} \end{vmatrix} \begin{vmatrix} I_0 \\ Q_0 \\ U_0 \\ V_0 \end{vmatrix} = L_S I_0 \quad (19)$$

where L_S is the magnetograph signal normalization matrix due to the effect of instrumental polarization.

6.2. Representation of the instrumental matrix of a solar telescope

In order to measure the telescope instrumental matrix, completely linearly and circularly polari-

zations are successively transmitted onto the telescope aperture and signals of all Stokes parameters are recorded in the continuum. With the purpose of accomplishing this procedure for a short time before and after taking magnetograms, a device is provided that consists of a polarizer and a phase plate, with an attached sun tracking drive (Grigoryev et al., 1980). The device is installed at the coelostat mirror mounting (Fig. 6) and can easily be introduced into and removed from the light beam incident on the coelostat mirror. The polarizer and the phase plate are of 300 mm in diameter, and the phase shift τ is $71^\circ.1$ for $\lambda = 5250 \text{ \AA}$. The incident unpolarized light $\{I, 0, 0, 0\}$ is transformed by a polarizer-phase plate system, introducing a phase shift τ :

$$\begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = M_z(\tau) R(\alpha) M_p(\chi) \begin{pmatrix} I \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (20)$$

where $M_p(\chi)$ is the polarizer matrix, χ is the angle of rotation of the entire system relative to the coordinate system of the magnetograph polarization analyzer, $R(\alpha)$ is the matrix of rotation of the phase plate with respect to the polarizer by an angle α , and $M_z(\tau)$ is the matrix of the phase plate.

1. The phase plate and polarizer are co-axial ($\alpha = 0^\circ$), and the system is rotated by a certain angle χ . The magnetograph signals will then be

$$\begin{aligned} S_I &= \frac{1}{2}(\alpha_{11} + \alpha_{12} \cos 2\chi - \alpha_{13} \sin 2\chi) I, \\ S_Q &= \frac{1}{2}(\alpha_{21} + \alpha_{22} \cos 2\chi - \alpha_{23} \sin 2\chi) I, \\ S_U &= \frac{1}{2}(\alpha_{31} + \alpha_{32} \cos 2\chi - \alpha_{33} \sin 2\chi) I. \end{aligned} \quad (21)$$

A series of measurements for $\chi = 11.45^\circ$ are performed in order to carry out an averaging of the results, this permitting the determination of the coefficients α_{K1} , α_{K2} , and α_{K3} for $K = 1, 2, 3$.

2. In each position of the system, prescribed by the angle $\chi = 11.45^\circ$, the phase plate is rotated with respect to the polarizer by the angles $\alpha = 45^\circ$ and $\alpha = -45^\circ$, and the magnetograph signals then are

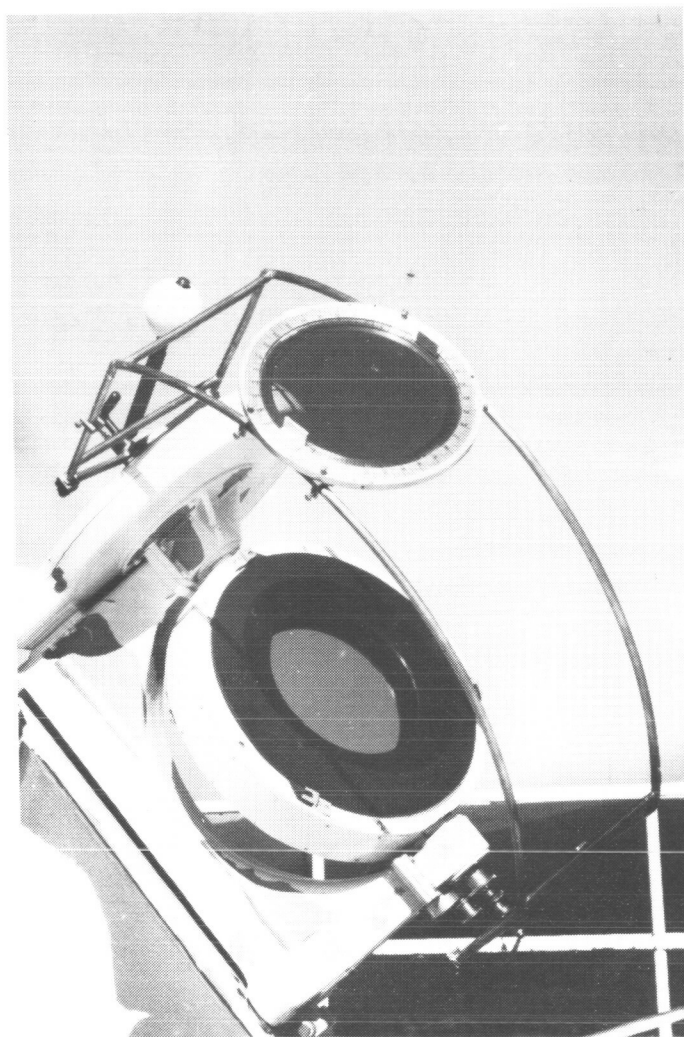


Fig. 6. Device for measuring the telescope
instrumental matrix

$$\begin{aligned}
S_Y &= (\alpha_{11} + \alpha_{12} \cos \tau \cos 2\chi - \alpha_{13} \cos \tau \sin 2\chi \pm \alpha_{14} \sin \tau) J, \\
S_Q &= (\alpha_{21} + \alpha_{22} \cos \tau \cos 2\chi - \alpha_{23} \cos \tau \sin 2\chi \pm \alpha_{24} \sin \tau) J, \\
S_U &= (\alpha_{31} + \alpha_{32} \cos \tau \cos 2\chi - \alpha_{33} \cos \tau \sin 2\chi \pm \alpha_{34} \sin \tau) J, \\
S_V &= (\alpha_{41} + \alpha_{42} \cos \tau \cos 2\chi - \alpha_{43} \cos \tau \sin 2\chi \pm \alpha_{44} \sin \tau) J,
\end{aligned} \tag{22}$$

where the plus sign before the last terms corresponds to $\alpha = 45^\circ$, and the minus sign to $\alpha = -45^\circ$. The expression (22) for different $\chi = \eta \cdot 45^\circ$ makes it possible to determine the coefficients α_{11} , $\alpha_{12} \cos \tau$, $\alpha_{13} \cos \tau$, and $\alpha_{14} \sin \tau$. Thus, the results of two successions of measurements make it possible to determine all the matrix coefficients as well as the phase shift τ of the plate, if not known before. The typical values of the coefficients α_{ij} of the matrix of signals L_{ij} , normalized to α_{11} , lie within the following limits:

$$\begin{vmatrix}
I & -0.0788 \pm 0.0461 & -0.0712 \pm 0.0014 & 0 \\
0.1078 \pm 0.2296 & 8.5707 \pm 9.9620 & -1.2764 \pm 0.6788 & 0 \\
-0.3291 \pm 0.2758 & 1.1871 \pm 2.1879 & 8.5819 \pm 9.5727 & 0 \\
0 & 0 & 0 & 8.5911 \pm 9.7403
\end{vmatrix} \tag{23}$$

6.3. Calibration of magnetograph channels

Measurements of Stokes parameters in solar magnetographs are essentially the relative measurements of the intensity in the sense that the magnitude of the parameters Q , U and V having the intensity dimensionality, is measured in units of the value of J , i.e. the total intensity of the radiation. Therefore, calibration of magnetograph channels is merely the determination of amplification coefficients of the channels or of the transfer function of the channels (of the electrooptical polarizer plus electron channels). If the electrooptical polarization analyzer is an ideal one, i.e. there is no cross-talk present in the Stokes parameter signals, then the transfer function may be written thus:

$$B = \begin{vmatrix}
\beta_{11} & 0 & 0 & 0 \\
0 & \beta_{22} & 0 & 0 \\
0 & 0 & \beta_{33} & 0 \\
0 & 0 & 0 & \beta_{44}
\end{vmatrix} \tag{24}$$

In the general case, due to the errors in the ad-

justment of the polarization analyzer, line miscentering and to errors in control voltage, the nondiagonal matrix elements may differ from 0. The transfer function matrix of the polarization analyzer and of electron channels can be represented independently by placing before the spectrograph slit a polarizer- $\lambda/4$ -plate system and by performing a series of measurements, similar to those of the telescope polarization matrix. It should be noted here, however, that in our description of the method for measuring the telescope polarization matrix we adopt a matrix $B = 1$. Indeed, the measuring procedure using a system consisting of a polarizer and a phase plate, determines a resulting matrix L_S , which involves a telescope polarization matrix L_T and a transfer function matrix of the analyzer and electron channels, B :

$$L_S = B \cdot L_T. \quad (25)$$

The diagonal matrix elements L_S (23) are virtually the amplification coefficients of electron channels with respect to the intensity channel.

7. Correction for Scattered Light

Scattered light in the Earth's atmosphere and within the instrument leads to a more than twofold decrease of magnetograph signals under average observing conditions (Bachmann et al., 1975). This appears to essentially account for the differences between the theoretical and empirical calibration curves, detected by Severny (1967).

In the reduction of magnetographic measurements we adopt a crude estimate for scattered light, ε , from the photosphere to the sunspot. The value of ε is derived by comparing the observed umbra-photosphere contrast with the true one (Selivanov, 1982):

$$\varepsilon = \frac{I'_u / I'_{ph} - 0.08}{0.92}, \quad (26)$$

where I'_u and I'_{ph} are the observed intensities in the umbra and photosphere, respectively. The correction of Stokes parameters measured in sunspots, S , is represented by:

$$S = S' \left(1 + \frac{\varepsilon}{1 - \varepsilon} \cdot 0.08 \right). \quad (27)$$

This reduction proves satisfactory for mean spatial resolution observations of large sunspots when intensity measurements in a sunspot umbra are little affected by the blurring effect.

In the near future it is planned to measure,

during magnetographic measurements, a total scattering and blurring function using a method described by Staveland (1972).

8. Handling of Measured Results

Processing of the measured magnetograph signals is carried out in the following order.

1. Normalization of the quantities S_0 , S_u and S_v to the quantity S_y yields a correction of signals for brightness variations with the result that values in units of the degree of polarization are obtained.

2. Reduction for the resulting instrumental matrix \mathbb{L}_S (19 and 23), defining the action of the telescope polarization matrix and of the transfer function of the channels, is performed by solving a set of four linear equations.

3. Reduction for scattered light of magnetograph signals in the sunspot region is performed using formula (27).

4. After the above cited corrections are introduced, a calculation is done of the magnetic field vector components. Since the computational procedure for inverse interpolation using two tables $f(H, r)$ and $\varphi(H, r)$ with two inputs in H and r is complicated and requires time-consuming computations, we have applied the separation of the variables H and r in the functions $f(H, r)$ and $\varphi(H, r)$. It is known that in the region of weak magnetic fields ($H \leq 500$ G), we can, to a high accuracy, employ the following approximation (see, e.g. Bachmann et al., 1975):

$$\begin{aligned} f &= f(H, r) \approx f(H, 0) \cos r, \\ \varphi &= \varphi(H, r) \approx \varphi(H, 90^\circ) \sin^2 r, \end{aligned} \quad (28)$$

and in the region of strong magnetic fields ($H > 500$ G) (Auer et al., 1977):

$$\frac{\varphi}{f} \approx \frac{\sin^2 r}{2 \cos r}. \quad (29)$$

We utilize an approximating expression which is, to a sufficient accuracy ($\Delta r \leq 5^\circ$), valid over a wide range of magnetic field vector strengths from 0 to 3000 G:

$$\tan^2 r \approx \frac{\varphi}{f} [0.5 + 2(\varphi + f)^2]. \quad (30)$$

Substitution into this expression for f and φ the measured values of signals S_v and $\sqrt{S_0^2 + S_u^2}$, respectively, yields a value of angle r . Using the tables of $f(H, r)$ and $\varphi(H, r)$, by inverse interpolation we derive the value of H .

9. Control of the Vector-Magnetograph

A solar image scanning program is specified either from the observer's console or from peripheral equipment. In the mode of peripheral control, both the direction along each coordinate and the scanning rate are specified by means of a code. The coordinate value in a binary code is put into peripheral equipment. Coordinates are read by means of 12-digit shaft encoders, connected to guide carriage lead screws.

Interfacing of magnetograph electron channels with the CM-2 computer (as with the USA HP-2116 computer) is accomplished via a trunk-module CAMAC system, placed directly at the telescope. Communication of the CAMAC equipment and a display module with the computer, that is at a distance of 150 m from the telescope, is accomplished via a two-cable line through an interface multiplex divider.

The computer software comprises two program packages. One comprises real-time tasks and provides collection of magnetograph data, as well as scanning and all system operation control. The observer-computer interaction is effected in the dialog mode by means of the display module. The observer calls in an appropriate program and loads numerical values of the parameters such as coordinates of the scanning area, scanning rate, etc. The computer executes the program, displays brief results on the display module and awaits further instructions. Source information is all recorded on magnetic disk.

The other program package comprises source information on processing programs, viz. reduction for instrumental polarization, scattered light, image quality, computation of heliographic coordinates of the scanned area, computation of magnetic field vector parameters and representation of information regarding the magnetic field and line-of-sight velocity on peripheral data display devices. Arrays of processed data together with reference data are transcribed onto magnetic tape for long-term storage.

To illustrate, Fig. 7 presents a chart of magnetic field distribution in a sunspot group of 5 October 1983 (E48 NO7). Heavy dotted lines show the location of sunspot's umbra and penumbra. Thin lines show isolines of the longitudinal component of the magnetic field H for the following values: 50, 100, 500, 1000 and 2000 G, where the solid lines denote N polarity and dotted lines denote S polarity. Arrows indicate the direction of the magnetic field vector in the plane of projection H , their length being proportional to the magnitude of H . The heavy solid line shows the location of the neutral line of longitudinal field H .

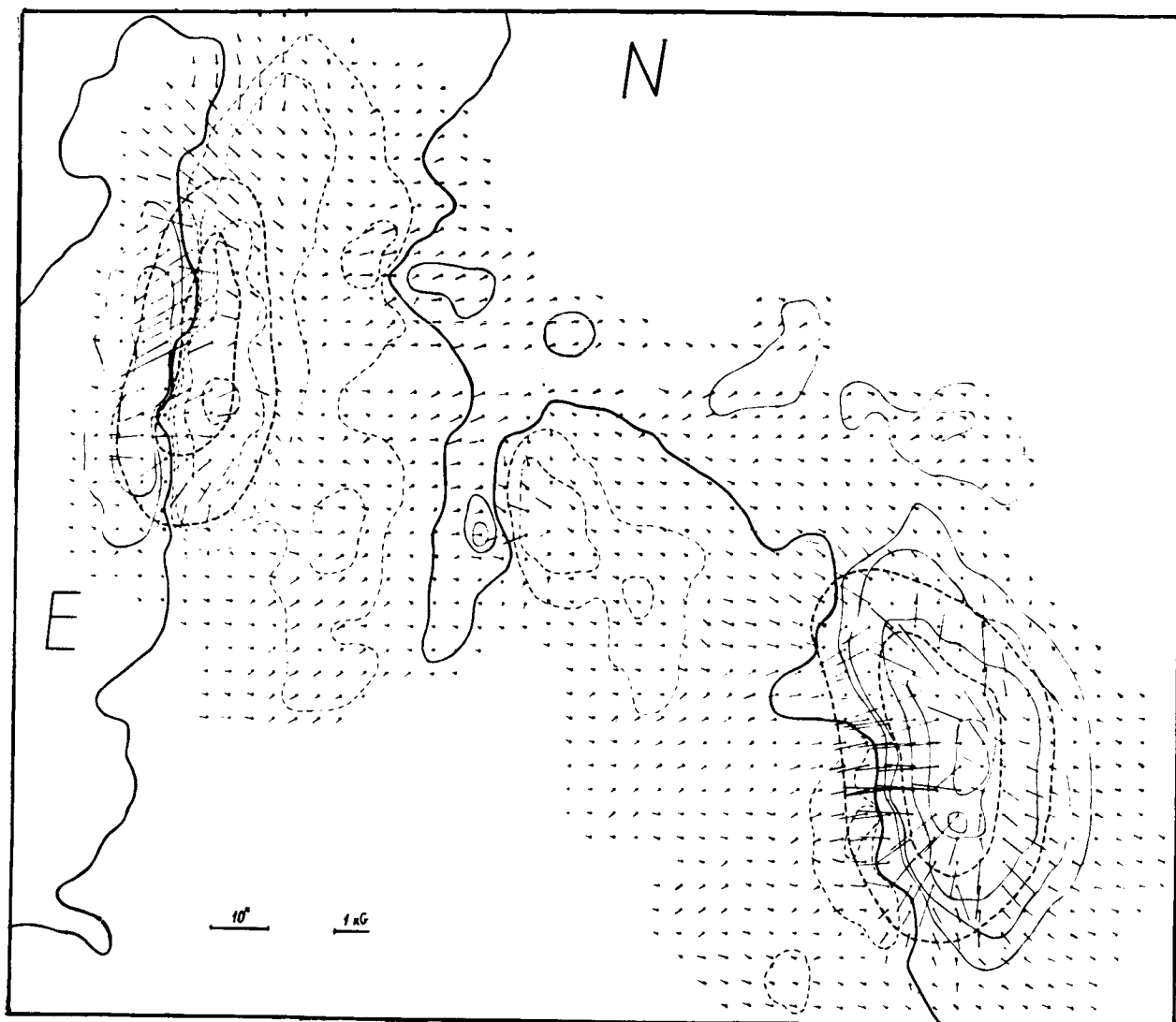


Fig. 7. An example of the magnetic field chart for the active region (see text).

Acknowledgements

We gratefully acknowledge the help of the many staff-members of the Astrophysical Instrument Making Laboratory of the SibIZMIR. We are also indebted to V.G.Mikhalevsky for his assistance in preparing the English version of the manuscript and for typing the text.

References

- Auer, L.H., Heasley, J.N. and House, L.L.: 1977 Solar Phys., 55, 47.
- Bachmann, G., Jager, F.W., Kunzel, H., Pflug, K. and Staude, J.: 1975, HHI-STP Report No. 4, Berlin.
- Duvall, T.U.: 1977, SUIPR Report No. 724, Stanford Univ.
- Grigoryev, V.M. and Katz, I.M.: 1975, Solar Phys., 42, 21.
- Grigoryev, V.M. and Selivanov, V.L.: 1978, Solnechnyye dannye, 2, 80.
- Grigoryev, V.M. and Kobanov, N.I.: 1977, Authors Certificate of the USSR No. 542161 "A method of controlling the magnitude of electrooptical modulator half-wave voltage". Bul. iz. No. 1.
- Grigoryev, V.M., Osak, B.F., Pflug, K. and Selivanov, V.L.: 1980, Phys. Sol.-Terr., 14, 81.
- Grigoryev, V.M. and Kobanov, N.I.: 1980, Phys. Sol.-Terr., 14, 77.
- Grigoryev, V.M. and Ilgamov, R.M.: 1983, Solnechnyye dannye, 8, 69.
- Jager, E.W.: 1972, Solar Phys., 27, 481.
- Ioshpa, B.A. and Obridko, V.N.: 1964, Geomagnetizm i aeronomiya, 4, 17.
- Katz, I.M.: 1973, in Issled. geomagn., aeron. i fiz. Solntsa, Moscow, Nauka, 28, 67.
- Kobanov, N.I.: 1974, Authors Certificate of the USSR No. 445154. Electrooptical switch control device. Bul. iz. No. 36.
- Kobanov, N.I.: 1979, Metrologiya, 10, 20.
- Kuznetsov, D.A., Kuklin, G.V. and Stepanov, V.E.: 1966, Results of the observations and investigations during the IQSY, Moscow, Nauka, 1, 80.
- Kotlyar, L.M.: 1961, Izv. Gl. astron. observ. v Pulkove, 167, 95.
- Livingston, W. and Harvey, J.: 1971, Contrib. Kitt Peak Nat. Obs. No. 558.
- Osak, B.F., Grigoryev, V.M., Kruglov, V.I. and Skomrovsky, V.I.: 1979, Novaya tekhnika v astronomii, Leningrad, Nauka, 84.
- Selivanov, V.L.: 1982, Issled. geomagn., aeron. i fiz. Solntsa, Moscow, Nauka, 60, 87.
- Severny, A.B.: 1967, Izvestiya Krymsk. astrofiz. observ. 36, 22.

- Shestov, N.S.: 1967, Separation of optical signals against random noise, Moscow, Sov. Radio.
- Shurcliff, W.A.: 1962, Polarized Light, Harvard Univ. Press.
- Stepanov, V.E., Grigoryev, V.M., Kobanov, N.I. and Osak, B.F.: 1975, Issled. geomagn., aeron. i fiz. Solntsa, 37, 147.
- Stepanov, V. E., and Severny, A. B.: 1962, Izv. Krimsk. Astrofiz. Obs., 28, 166.

MULTICHANNEL BIREFRINGENT FILTER

AI GOUXIANG and HU YUEFENG

Beijing Observatory, Academia Sinica

Our birefringent filter has a large field-of-view and no additional polarization. It plays an important role in observing the solar monochromatic image and the solar vector magnetic field (Beckers, 1975; Hagyard et al., 1982; Ai et al., 1984). But, at present, it has only one channel. For simultaneous multichannel observations, the solar spectrograph is better than the birefringent filter. A suggestion has been proposed to try to obtain a multichannel birefringent filter which will be used in a new telescope at the Huairou reservoir station of Beijing Observatory.

In the single channel birefringent filter, a considerable amount of light is absorbed by the polaroids. If we use polarizing beam splitters (Chapman et al., 1969) instead of the polaroids, two spectral components will emerge, polarized at right angle to each other, and two independent filter channels are obtained. (Fig. 1)

By means of N polarizing beam splitters, $(N+1)$ channels can be divided. In principle, any number of limitless channels can be obtained, thereby subdividing the whole solar spectrum. But since the space in a telescope is limited the channels to be used are also limited. For example, its numbers will be 3, 5, 9, or 25, etc.

For the new telescope, 5 and 9 channels are being considered, and the spectral range is from λ 3800Å to λ 7000Å. Many lines are included in this range, for example, H, K, H_β , $\lambda\lambda$ 5324Å, 5250Å, 6302Å, $H\alpha$, etc., and some of the lines are suited to measure solar velocity fields. According to the character of these lines, the half width of each channel is determined. Moreover, in some channels the solid polarizing Michelson interferometer is considered for measuring velocity field with a 1m/s accuracy.

The multichannel birefringent filter is characterized by the following advantages:

1. Solar monochromatic image, solar magnetic field and velocity field in different lines can be simultaneously obtained in a simple telescope. This is a special advantage for space solar observatories, big solar telescopes and telescopes of single stations. The efficiency of observations will be greatly increased.

2. The range of working wavelength in each channel is 500Å - 1000Å, while in universal filters (Beckers, 1975), it is 3000Å. Hence in the multichannel filter the performance of an achromatic wave plate is tolerated.

H and K lines which are not included in the working range of the universal filter will be included in one of the channels of the multichannel filter.

3. Various suitable band widths can be selected to match each line separately.

4. Using a single telescope under the same atmospheric conditions, various solar monochromatic images, magnetic field maps and velocity field maps can be obtained simultaneously, which will promote research of the solar activity phenomena.

The following problems remain to us to be solved:

1. Because the polarizing efficiencies of polarizing beam splitters are about 0.99, a polaroid with polarizing efficiency about 0.90 is necessary to supplement each polarizing beam splitter. This kind of polaroid will be manufactured specially.
2. Because the transmission wavelength of each element is tunable, difficulties may occur in assemblage, adjustment and control must be overcome.
3. Difficulties with thermostats caused by the complex structure must be resolved.
4. The selection of the working wavelength may cause complexities in the optical design.

At the present, the design is being prepared.

References

- Beckers, J. M. (1975): Appl. Opt., 14, 2061.
- Hagyard, M. J., Cumings, N. P., West, E. A., and Smith, J. E. (1982): Solar Phys., 80, 33.
- Ai Guoxiang, Hu Yuefeng, et al. (1984): The Birefringent Filter for Measurement of Both the Solar Magnetic Vector Field and Line of Sight Velocity Field, in proc. Kunming Workshop On Solar Physics and Interplanetary Travelling Phenomena.
- Chapman, P. B., Downs, M. J., and King, R. T. (1969): Appl. Opt., 8, 1965.



SECTION 3

**TECHNIQUES FOR INTERPRETATION OF
OBSERVATIONAL DATA**

DIAGNOSTICS OF VECTOR MAGNETIC FIELDS

J.O. STENFLO

Institute of Astronomy, ETH-Zentrum, CH-8092 Zürich, Switzerland

Abstract. If we disregard coherence effects (Hanle effect), the theory of line formation in a magnetic field is reasonably well developed and can be applied to any magnetic-field configuration and model atmosphere, including non-LTE and magneto-optical effects. Nevertheless there are fundamental problems in the determination of the vector magnetic field, not so much because of instrumental problems or lack of theoretical tools, but because of the spatially unresolved fine structure of the magnetic field. Whereas the circular polarization responds to magnetic flux, the linear polarization represents something that is more related to average magnetic energy. Previous interpretations, which have been based on the assumption of a spatially resolved field, may therefore result in greatly distorted vector fields (generally with much too strong apparent horizontal fields), since the spatial averaging affects the longitudinal and transverse Zeeman effects in so fundamentally different ways.

These and other diagnostic problems of fluxtube physics are reviewed, and it is indicated how they may be resolved by using extended spectral information.

1. Introduction

The systematic mapping of vector magnetic fields on the sun was pioneered by Stepanov and Severny (1962) at Crimea, who presented maps of the vector field recorded as early as 1959, and gave a thorough treatment of the analysis procedure. The technique was essentially a Babcock-type photoelectric magnetograph, extended to record simultaneously not only the circular polarization, but the linear polarization and its orientation as well. Stepanov and Severny (1962) realized that magnetooptical effects influenced the determination of the field azimuth if the exit slits were close to the line center, but that these effects could be avoided by using the outer line wings for the diagnostics. Rachkovsky (1962a,b) had developed the theory of magnetooptical effects in the radiative transfer of the Stokes vector, and this is the theory that is used today.

Using the Crimean magnetograph, spiralling magnetic patterns around sunspots were found (Stepanov and Severny, 1962) as well as a rotation with depth of the field azimuth (Severny, 1964, 1965), determined by observing in different portions of the line profile and in lines formed at different depths.

Kotov (1972) extended the technique of using vector-field recordings to determine the electromagnetic structure of a unipolar sunspot. The Maxwell equation $\mathbf{j} = \frac{c}{4\pi} \nabla \times \mathbf{H}$ was used to derive not only the vertical current density j_z from the horizontal gradients of the transverse field, but also to derive the horizontal current density using the vertical gradients obtained from observations in lines formed at different heights. The height gradient of the longitudinal field $\partial B_z / \partial z$ was obtained from the $\text{div } \mathbf{B} = 0$ condition.

Since these earlier investigations, several groups have made efforts to record the vector magnetic field, most recently Baur et al. (1980, 1981) by recording the line profiles of the Stokes vector, and Hagyard et al. (1982), using a narrow-band filter to produce monochromatic images of the four Stokes parameters.

Although many maps of vector magnetic fields have thus been produced in the past, the conclusions derived from these maps should be treated with great caution, since the conversion of the polarization data into magnetic fields has been based on the assumption that the field has been spatially resolved. In the case of the longitudinal component of the magnetic field there is no big problem, since the amount of circular polarization is roughly proportional to the longitudinal magnetic flux. The whole problem arises when trying to combine the circular and linear polarization measurements to obtain a vector magnetic field (averaged over the resolution element). The reason for the trouble is that the linear polarization does not scale at all like transverse magnetic flux, but more like transverse magnetic energy. Thus as the spatial averaging works in so fundamentally different ways for the longitudinal and transverse Zeeman effects, the vector magnetic field cannot be constructed unless one has an adequate model of what the unresolved structure looks like. Although it is well established that the flux outside sunspots consists of subarcsec kG fields (Stenflo, 1973, 1976; Harvey, 1977), and although empirical fluxtube models are available (Stenflo, 1975; Chapman, 1977, 1979), these models are hardly accurate enough to permit a reliable determination of the vector field.

In the present paper we will review the main problems in diagnosing the vector magnetic field, and indicate how they can be successfully solved by using extended spectral information.

2. General Properties of Stokes Profiles

The equation of transfer of polarized light can be written

$$\mu \frac{d\mathbf{I}}{d\tau} = \eta \mathbf{I} - \mathbf{S}. \quad (2.1)$$

\mathbf{I} is the Stokes vector containing the standard four Stokes parameters I , Q , U , V . μ is the cosine of the heliocentric angle, τ the optical depth, η the absorption matrix, and \mathbf{S} the source function vector. In LTE, without scattering, $\mathbf{S} = \eta \mathbf{B}$, \mathbf{B} being the Planck function times the unit vector.

The form of η was first derived by Unno (1956) and extended by Rachkovsky (1962a) to include magneto-optical effects. This Unno-Rachkovsky formula forms the basis of all contemporary calculations of transfer of polarized radiation in solar spectral lines. It is fairly straightforward to extend it to the non-LTE case, as has been done in a general way by House and Steinitz (1975), and by Stenholm and Stenflo (1978) for multi-dimensional non-LTE radiative transfer in fluxtube geometries.

The Unno-Rachkovsky formulation uses a linear polarization basis for the definition of the Stokes parameters. This allows arbitrary geometries and line-of-sight variations of the physical parameters to be treated, but there is strong cross-talk in the radiative-transfer process between the various Stokes parameters (the absorption matrix η has large off-diagonal elements). The formulation developed independently by Stepanov (1958a,b) uses a basis of mutually orthogonal elliptical polarization such that the absorption matrix (in the absence of magneto-optical effects) becomes diagonal. The four transfer equations of the Stokes parameters then decouple from each other and can be solved independently by standard methods for unpolarized radiation. The disadvantage of this formulation is that it cannot handle arbitrary geometries of the magnetic field. The relation between the Unno-Rachkovsky and the Stepanov formulations is explained in Stenflo (1971).

For a longitudinal magnetic field, the orthogonal polarization basis in the Stepanov formulation becomes the left and right handed circular polarizations of the σ components, for a transverse field it becomes the two linear polarizations of the σ and π components. Since the σ and the π components (disregarding magneto-optical effects) decouple in the radiative transfer process, we can express the emergent I , Q , U , and V in terms of the emergent $I_{\sigma 1,2}$ and I_{π} , which can be expressed as the Zeeman-shifted

I profile. The polarization Q , U , and V can thus be expressed in terms of I regardless of the atmospheric structure.

As the σ components are circularly polarized for a longitudinal field,

$$I_{\sigma 1,2} = \frac{1}{2}(I \pm V). \quad (2.2)$$

Assuming no height variation of the field strength B , we can write

$$I_{\sigma 1,2}(\lambda) = \frac{1}{2}(I \pm \Delta\lambda_H), \quad (2.3)$$

where the Zeeman splitting $\Delta\lambda_H$ is

$$\Delta\lambda_H = 4.67 \times 10^{-13} g \lambda^2 B. \quad (2.4)$$

g is the Landé factor, B should be given in G, and the wavelength in Å. (2.2) and (2.3) give

$$V = \frac{1}{2}[I(\lambda + \Delta\lambda_H) - I(\lambda - \Delta\lambda_H)]. \quad (2.5)$$

If $\Delta\lambda_H$ is small compared with the line width, only the first term in the Taylor expansion

$$V = \Delta\lambda_H \frac{\partial I}{\partial \lambda} + \dots \quad (2.6)$$

needs to be retained.

Fig. 1 demonstrates that (2.6) is indeed a valid relation for many strong lines, for which the weak-field approximation applies (since the line width is large). The discrepancies between the V and $\partial I/\partial \lambda$ curves in the case of the Na I D₁ line are caused by the telluric contaminations of the I spectrum. In the case of the H β line the deviations occur in the solar blend lines, since their Landé factors are different

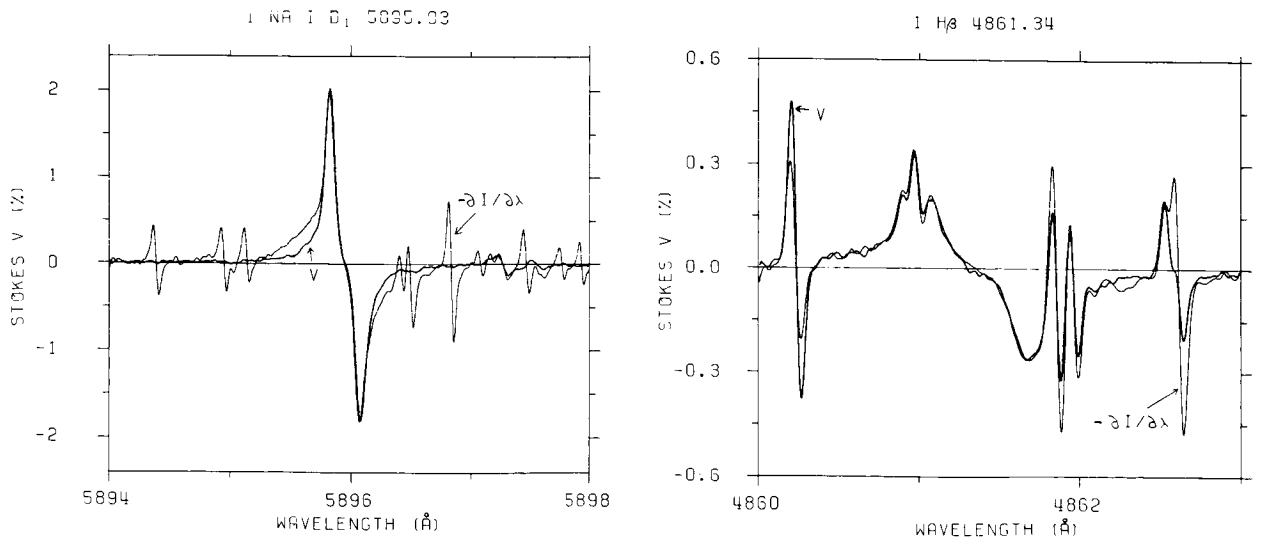


Fig. 1. Stokes V (thick curve) and intensity gradient $\partial I/\partial \lambda$ (thin curve) around the Na I D₁ line (left diagram) and the H β line (right diagram). $\partial I/\partial \lambda$ has been multiplied by 4.3 mÅ (1.8 mÅ) in the left (right) diagram, to normalize its amplitude to that of Stokes V . The many telluric lines of water vapor around the Na I D₁ line are unpolarized.

from that of the $H\beta$ line, and the normalization of $\partial I/\partial \lambda$ has been made with respect to the V amplitude in $H\beta$.

Thus in spite of the circumstance that LTE theory has no application whatsoever to the Na I D_1 and $H\beta$ lines, the simple relations (2.5) and (2.6) apply due to the diagonalization of the absorption matrix.

Similarly for a transverse magnetic field, we get

$$I_{\pi,\sigma} = \frac{1}{2}(I \pm Q), \quad (2.7)$$

assuming that the positive Q direction is defined to be parallel to the magnetic field.

$$\begin{aligned} I_{\pi} &= \frac{1}{2}I(\lambda), \\ I_{\sigma} &= \frac{1}{4}[I(\lambda + \Delta\lambda_H) + I(\lambda - \Delta\lambda_H)]. \end{aligned} \quad (2.8)$$

From (2.7) and (2.8) we obtain

$$Q = \frac{1}{2}\{I(\lambda) - \frac{1}{2}[I(\lambda + \Delta\lambda_H) + I(\lambda - \Delta\lambda_H)]\}. \quad (2.9)$$

The Taylor expansion gives

$$Q = -\frac{1}{4}(\Delta\lambda_H)^2 \frac{\partial^2 I}{\partial \lambda^2} + \dots \quad (2.10)$$

In the weak-field approximation only the first term has to be considered.

As seen from (2.8), (2.10), and (2.4), $V \sim B$ and $Q \sim B^2$ for weak fields. Further, because the absorption coefficients for Q and V are proportional to $\sin^2 \gamma$ and $\cos \gamma$, respectively, γ being the angle between the field vector and the line of sight, we can write for any field strength B :

$$\begin{aligned} Q &= q(B) \sin^2 \gamma, \\ V &= v(B) \cos \gamma, \end{aligned} \quad (2.11)$$

where $q(B) \sim B^2$ and $v(B) \sim B$ for weak fields. For strong fields (complete Zeeman splitting), $q(B)$ and $v(B)$ become comparable in magnitude, but for weaker fields, $q(B) \ll v(B)$. This makes it much harder to measure the transverse field as compared with the longitudinal field. q and v have a slight dependence on γ , which can generally be neglected.

(2.11) and the weak-field approximation show that while the circular polarization is a measure of the longitudinal magnetic flux, the linear polarization is more a measure of the transverse magnetic energy as averaged over the resolution element. Because of the fundamentally different averaging processes of the spatially unresolved magnetic fine structures, the derived vector magnetic field will be severely distorted unless the unresolved structures can be properly accounted for. This is the reason why vector magnetic fields have not yet been reliably determined outside sunspots. This problem will be treated in the next section.

The analytic solution of the transfer equations in the case of a Milne-Eddington atmosphere with a homogeneous magnetic field and a Planck function with linear optical-depth dependence, including magneto-optical effects, was first found by Rachkovsky (1962b) (cf. Stenflo, 1971). The temperature structure of the atmosphere, or the line strength, can be conveniently characterized by one parameter η_0 , the ratio between the line absorption coefficient at line center and the continuous absorption coefficient.

In Fig. 2 we have used the Rachkovsky solution to plot the four Stokes profiles for different field strengths B , assuming $\gamma = 45^\circ$, χ (azimuth of B) = 0, and $\eta_0 = 1.72$, which reproduces the depth of

$$\gamma = 45^\circ, \chi = 0$$

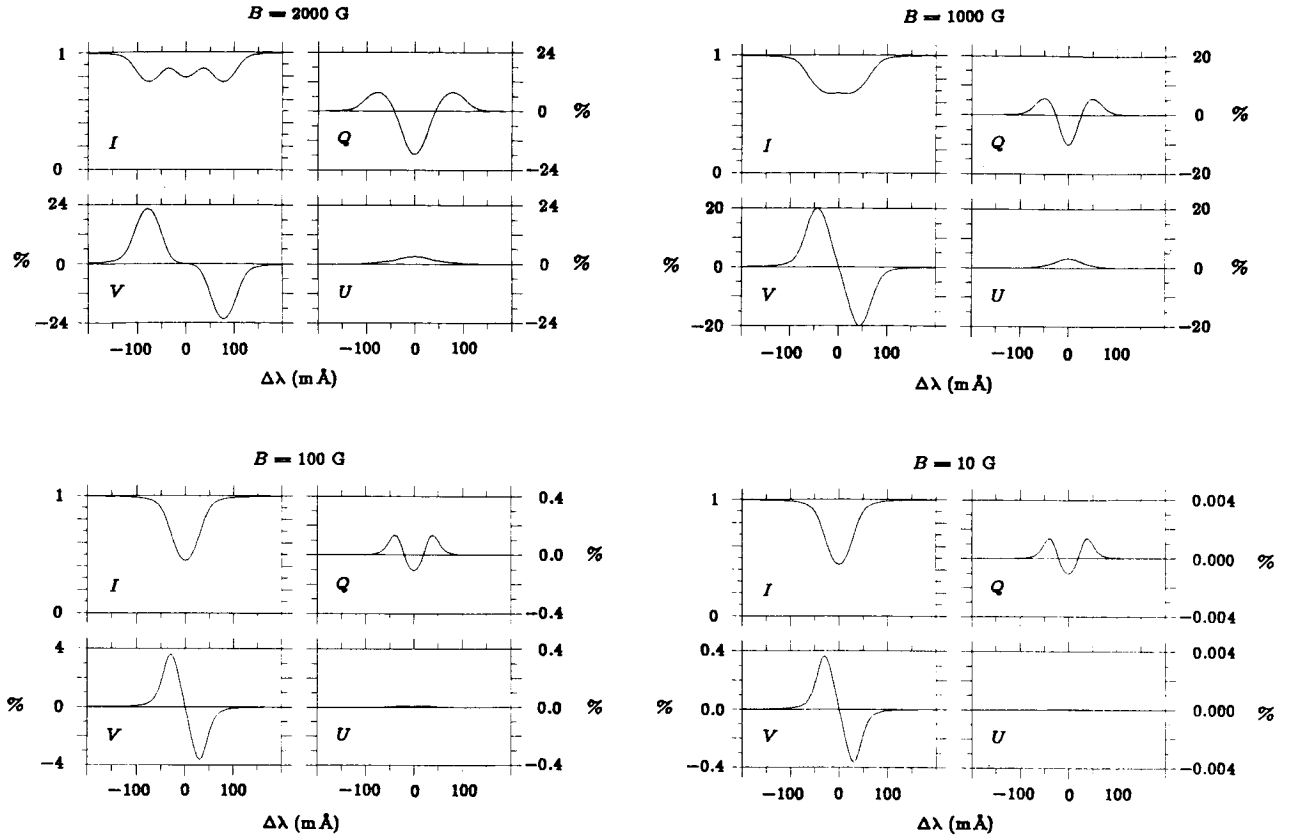


Fig. 2. Stokes line profiles, calculated with a Milne-Eddington model atmosphere including magneto-optical effects, assuming a field inclination $\gamma = 45^\circ$, field azimuth $\chi = 0$, $\lambda = 5250.22 \text{ \AA}$, Landé factor $g = 3$, Doppler width 25.25 m\AA , damping parameter 0.25 , and line-strength parameter $\eta_0 = 1.72$. Q , U , and V are given in units of the intensity of the adjacent continuous spectrum.

the Fe I $\lambda 5250.22 \text{ \AA}$ line in the fluxtube model of Stenflo (1975). The damping constant is assumed to be 0.25 , and the Doppler width 25.25 m\AA (to obtain the same line width as found in our Fourier transform spectrometer data (Stenflo et al., 1984).

In the absence of magneto-optical effects U would be exactly zero (since $\chi = 0$). For strong fields, however, a substantial U signal appears at the line center due to magneto-optical effects, but this phenomenon disappears for weaker fields. The main importance of the magneto-optical effects is with respect to determinations of the field azimuth χ when the central portion of the line is used. The determination of γ and B is however hardly affected at all. Since the magneto-optical effects have long been well understood, there is no reason not to include them in the computer calculations apart from saving some computer time. There is no new free parameter that has to be introduced when including these effects, since the damping constant and Doppler width of the anomalous dispersion function are the same as already used in the Voigt function.

Fig. 2 also demonstrates how the ratio between the Q and V amplitudes rapidly goes to zero when B becomes small, due to the B^2 dependence of Q for weak fields. For strong fields Q and V are comparable in magnitude. In the diagrams for $B = 2000$ and 1000 G in Fig. 2, the polarization scales for Q , U , and

V are identical. In the $B = 100$ G diagram, we have reduced the Q and U scales by a factor of 10 with respect to V , in the $B = 10$ G diagram they have been reduced by a factor of 100. The profiles in the $B = 100$ and 10 G diagrams look identical, but if we compare the polarization scales used, we see that V has scaled as B , whereas Q has scaled as B^2 , as expected from our Taylor expansion treatment.

To get an impression of the applicability of the B and B^2 scaling, we have calculated $v(B)$ (as V for $\gamma = 0$) and $q(B)$ (as Q for $\gamma = \pi/2$) in (2.11) for the same Milne-Eddington parameters as used for Fig. 2. The results of course depend on where in the line profile one observes. To illustrate the effects, we have assumed that a narrow-band filter with a full width at half maximum of $1/8 \text{ \AA}$ is used. A $(\sin x/x)^2$ transmission profile is assumed, but the sidebands are disregarded. The calculations have been done for two different positions of the filter passband, at $\Delta\lambda = 40$ and 80 m\AA from line center, respectively. The 40 m\AA position is close to where the V profile has a maximum, and is just outside the zero crossing of the Q profile, as seen in Fig. 2. If one goes closer to the line center the V signal goes down, magnetooptical effects become appreciable and the Q zero crossing becomes disturbing. Out at $\Delta\lambda = 80 \text{ m\AA}$ one avoids magnetooptical effects and the Q zero crossing, but the signals are smaller.

In Fig. 3 we have plotted $v(B)/B$, normalized to unity for $B = 0$. The deviation from the linear B dependence of $v(B)$ gradually increases with field strength and is larger for the filter position closer to the line center, as expected. A vertical dashed line is drawn at $B = 1 \text{ kG}$, the typical field strength in plages as well as in the quiet network. At this value of B , the deviation from the linear dependence of V on B is only 10 - 15 %. This applies to the 5250 \AA line, which has a Landé factor of 3. For practically any other line in the solar spectrum, the deviation from linearity will be smaller.

In the case of Q , the deviation from the B^2 dependence becomes large already for relatively weak fields, as shown by the right diagram of Fig. 3, where we have plotted $q(B)/B^2$, normalized to unity for

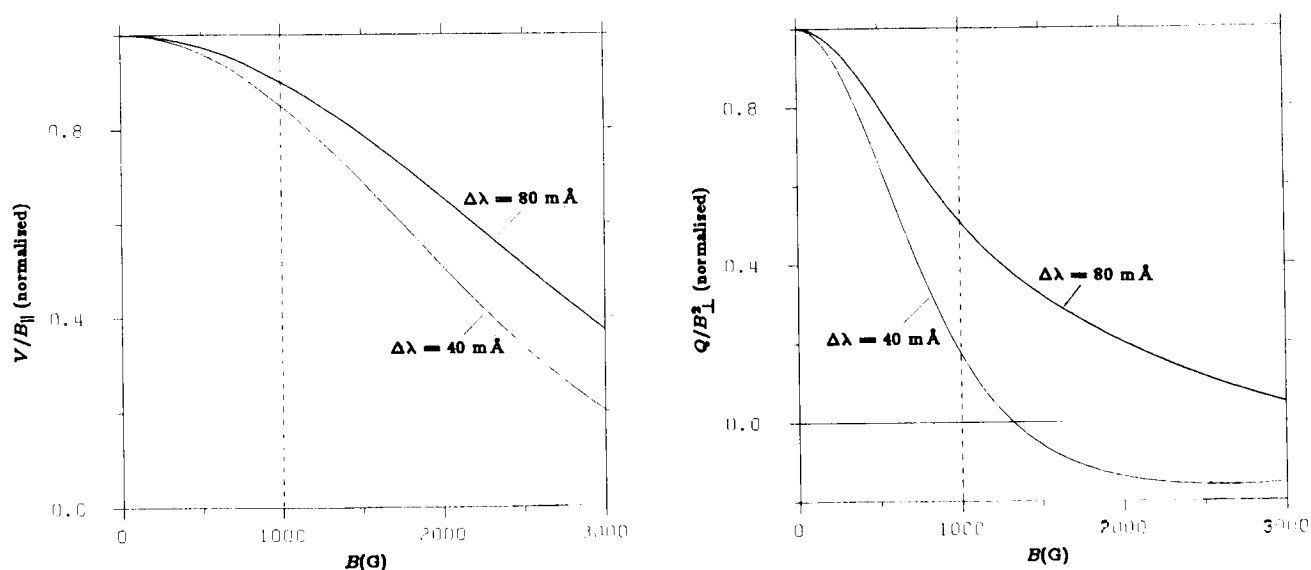


Fig. 3. Left diagram: Stokes V/B vs. field strength B , assuming a longitudinal field. Right diagram: Stokes Q/B^2 vs. B , assuming a transverse magnetic field. The calculations are based on the same parameters as used for Fig. 2, but it is assumed that the Stokes V and Q line profiles are sampled by a $1/8 \text{ \AA}$ filter positioned at $\Delta\lambda = 40 \text{ m\AA}$ (thin line) and 80 m\AA (thick line) from line center. The curves have been normalized to unity for $B = 0$. The vertical dashed line indicates the typical fluxtube field strength (1 kG).

$B = 0$. The curve for the 40 mÅ filter position has a zero crossing at $B = 1.3$ kG, since for strong fields the zero-crossing point in the Q profile is moved out beyond 40 mÅ, as shown by Fig. 2. These functions will be used in the next section for the calculation of the effect of the spatially unresolved magnetic field on the interpretation of vector magnetic field measurements.

3. Influence of the Spatially Unresolved Magnetic Structure

It is well established that practically all magnetic flux outside sunspots occurs in strong-field (1 - 2 kG) form (cf. Stenflo, 1976; Harvey, 1977). For the interpretation of measurements of the longitudinal magnetic field, a two-component model has proven to be very useful (Stenflo, 1973; Frazier and Stenflo, 1978). It is convenient to use such a model also to estimate the influence of the spatially unresolved magnetic structures on the interpretation of measurements of the vector magnetic field.

Thus, regardless of whether our spatial resolution element is $0.5''$, $5''$, or any other value, we assume that the fraction α of this resolution element is covered by magnetic fields of strength B , inclination angle to the line of sight γ , and azimuth angle χ . The remaining fraction $1 - \alpha$ is assumed to be field free. Whereas the line strength in the non-magnetic atmosphere is characterized by Milne-Eddington parameter η_0 (cf. Section 2), the corresponding parameter in the magnetic atmosphere is η_m , which takes care of the temperature line weakenings in the magnetic regions. For simplicity we assume the Doppler width and damping parameter to be the same in both the magnetic and non-magnetic component.

If we assume that we have chosen the coordinate system for the Stokes parameters such that the positive Q direction is parallel to the magnetic field, we do not need to consider the azimuth angle, and can use the definitions of q and v in (2.11) to obtain

$$\begin{aligned} I &= \alpha I(B, \eta_m) + (1 - \alpha) I(0, \eta_0), \\ Q &= \alpha q(B, \eta_m) \sin^2 \gamma, \\ V &= \alpha v(B, \eta_m) \cos \gamma. \end{aligned} \tag{3.1}$$

Here we have made the dependence of Q and V on temperature effects explicit via η_m in q and v , but we disregard the minor variation of q and v with γ .

When the vector magnetic field is recorded with a filter magnetograph, it is the four Stokes parameters I , Q , U , and V that are observed, with α , η_m , B , γ , and χ as the five unknown parameters. By choosing the orientation of our Stokes system such that $\chi = 0$, we have eliminated the U information, such that we now are considering four unknown parameters and three input parameters. If $\alpha \ll 1$, the I signal in (3.1) contains practically no information on α and η_m , so this leaves us with two observables (Q and V) and four unknowns (α , η_m , B , and γ). Clearly such observational input is not sufficient for a determination of the vector field (B and γ). The observational basis needed is Stokes line profiles with very high spectral resolution and/or simultaneous recordings in carefully selected combinations of spectral lines.

In all previous observational work on vector magnetic fields the spatially unresolved structure has been ignored. With our two-component model, the disregard of the unresolved structure corresponds to assuming $\alpha = 1$ and $\eta_m = \eta_0$. This leaves us with the two unknown parameters B and γ , which can be determined from the two observables Q and V . As however α is generally < 1 , and $\eta_m \neq \eta_0$, the values derived for B and γ do not represent the actual field strength and inclination, but some apparent vector field, which we will distinguish from the actual field by using index *app*. Accordingly

$$Q = q(B_{app}, \eta_0) \sin^2 \gamma_{app}, \quad (3.2)$$

$$V = v(B_{app}, \eta_0) \cos \gamma_{app}.$$

We will now calculate the relation between B_{app} , γ_{app} and B , γ , to estimate how large the errors in the derived values will be.

B_{app} can be derived from Q and V using the condition

$$\frac{Q}{q(B_{app}, \eta_0)} + \left[\frac{V}{v(B_{app}, \eta_0)} \right]^2 = 1, \quad (3.3)$$

which follows from (3.2). Knowing B_{app} , we find γ_{app} from the relation

$$\frac{\sin^2 \gamma_{app}}{\cos \gamma_{app}} = \frac{Q v(B_{app}, \eta_0)}{V q(B_{app}, \eta_0)}, \quad (3.4)$$

which also follows from (3.2).

For a given field structure, i.e., for given values of α , η_m , B , γ , we can calculate Q and V from (3.1). These are the "observed" Q and V to be inserted in (3.3) and (3.4), from which B_{app} and γ_{app} are calculated. Thus we find the dependence of the apparent field on the four model parameters.

Unless otherwise stated, the following diagrams will be based on the same 5250.22 Å line parameters as in Fig. 3, using a 1/8 Å filter passband centered at $\Delta\lambda = 40$ and 80 mÅ, respectively. $\eta_0 = 4.4$, which reproduces the relative line depth of 0.74 found for the 5250 Å line in our FTS recordings in non-magnetic regions. For η_m we use either 1.72, which reproduces the temperature line weakening in the fluxtube model

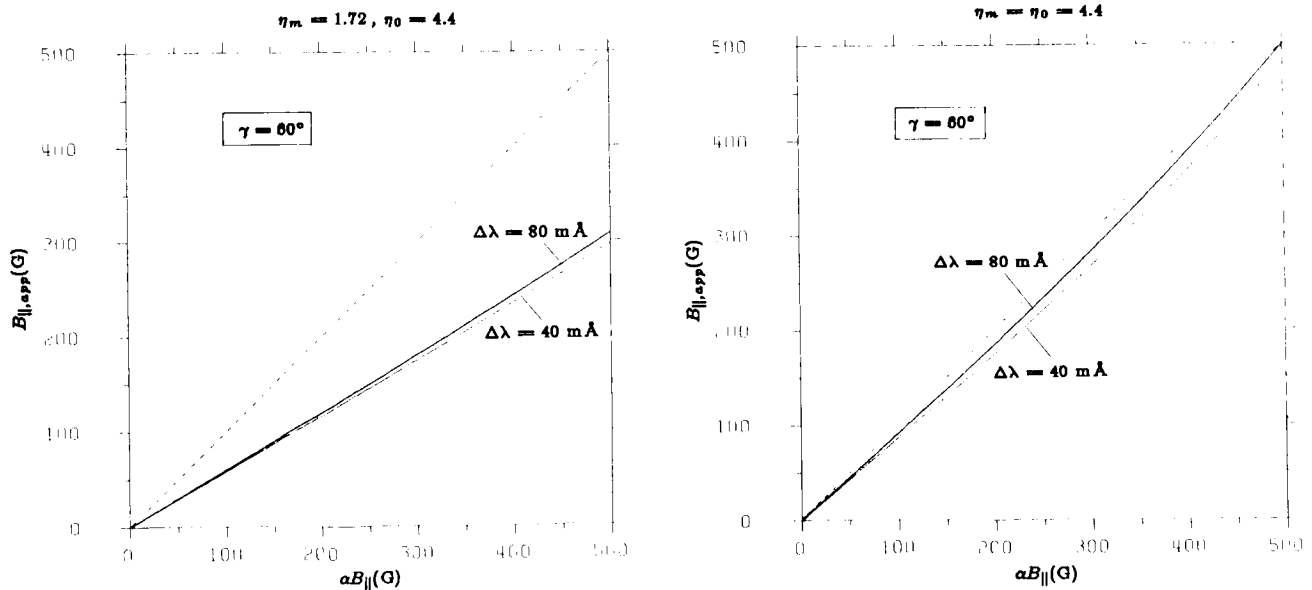


Fig. 4. Apparent longitudinal magnetic field vs. true average longitudinal field, derived from a two-component model with magnetic area factor α and intrinsic field strength $B = 1$ kG, for the same filter positions and line parameters as in Fig. 3. The case when the field inclination angle $\gamma = 60^\circ$ is illustrated. Left diagram: Temperature line weakenings are included (via the use of $\eta_m = 1.72$). Right diagram: No line weakenings. The dashed 45° line indicates the desired relation for correct diagnostics.

of Stenflo (1975), or 4.4, to separate the effects caused by η_m and α . Further, we assume that the true field strength B is always 1000 G, the value obtained for a two-component model with rectangular fluxtube cross section, valid for both network and plages (Stenflo, 1973; Frazier and Stenflo, 1978).

Fig. 4 gives the relation between the apparent longitudinal magnetic field, $B_{\parallel,app} = B_{app} \cos \gamma_{app}$, and the true average longitudinal field, $\alpha B_{\parallel} = \alpha B \cos \gamma$, for an inclination $\gamma = 60^\circ$ of the field (i.e., $\alpha B_{\parallel} = 500\alpha$ G). For a correct diagnostics of the longitudinal field, the curves should fall along the dashed 45° line. A comparison between the left and right diagrams of Fig. 4 shows that the deviation from the 45° line is almost entirely caused by the temperature weakening of the line. This causes the average apparent longitudinal component to be too small by about 40 % in our model. The diagrams for other values of γ look practically identical (except that the values on the axes scale with $\cos \gamma$).

In Fig. 5 we have plotted the apparent transverse magnetic field, $B_{\perp,app} = B_{app} \sin \gamma_{app}$, versus the true average transverse field, $\alpha B_{\perp} = \alpha B \sin \gamma$, for an inclination $\gamma = 30^\circ$. The results look the same for other inclinations, if we scale the values on the axes with $\sin \gamma$. The deviation from the correct 45° line is much larger than for the longitudinal field, and it is sensitive to where in the line profile that the passband is located. A comparison between the left and right diagrams in Fig. 5 further shows that the discrepancy is not only due to the line weakening; it is strongly related to the value of α . For small values of α the slope of the curves goes to infinity instead of to unity.

The error that the unresolved fine structure introduces becomes even more striking when considering the relation between the apparent and true field inclinations, as illustrated in Figs. 6 and 7. Even for $\alpha = 1$ a considerable deviation from the 45° line occurs for filter position $\Delta\lambda = 40$ mÅ, due to the line weakening (this discrepancy is removed when $\eta_m = \eta_0$). For small values of α the apparent inclination γ_{app} becomes much larger than the true field inclination γ . As seen by Fig. 7, this is almost entirely an effect of α . The main effect of the line weakening is to lower the curve for $\Delta\lambda = 40$ mÅ somewhat. When α decreases, the apparent field looks more and more transverse. For vanishing values of α , $\gamma_{app} = 90^\circ$ regardless of the value of γ (if we exclude the immediate neighbourhood of $\gamma = 0$). This effect causes the

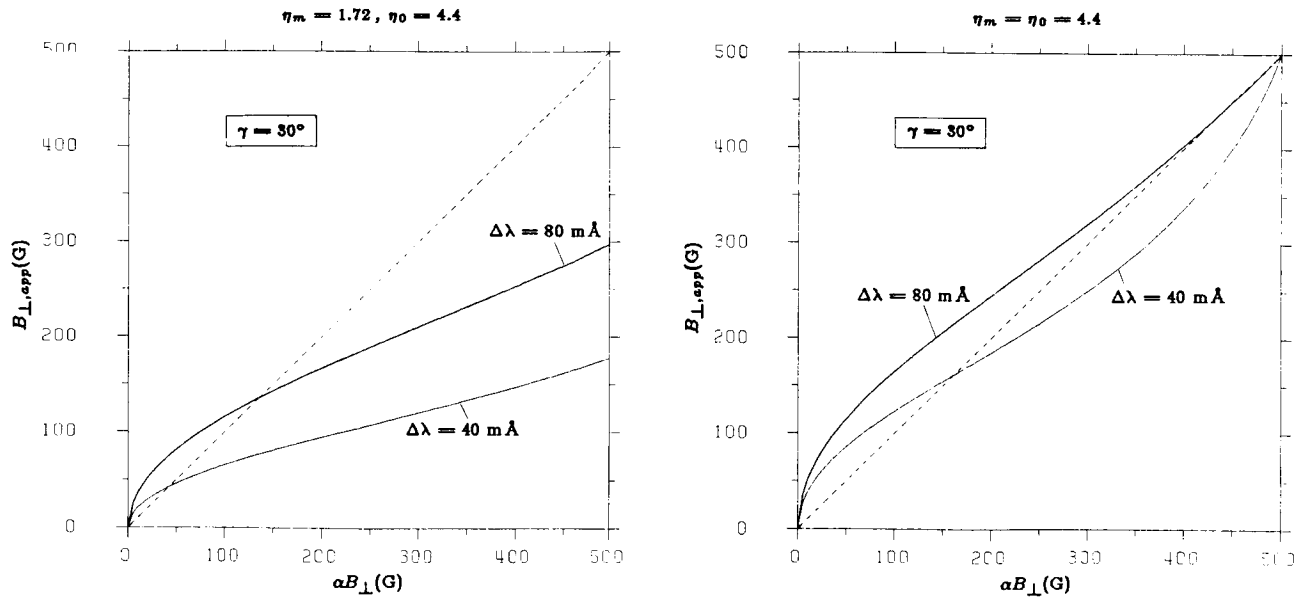


Fig. 5. Apparent transverse magnetic field vs. true average transverse field, for the case of $\gamma = 30^\circ$. Left (right) diagram: With (without) line weakenings.

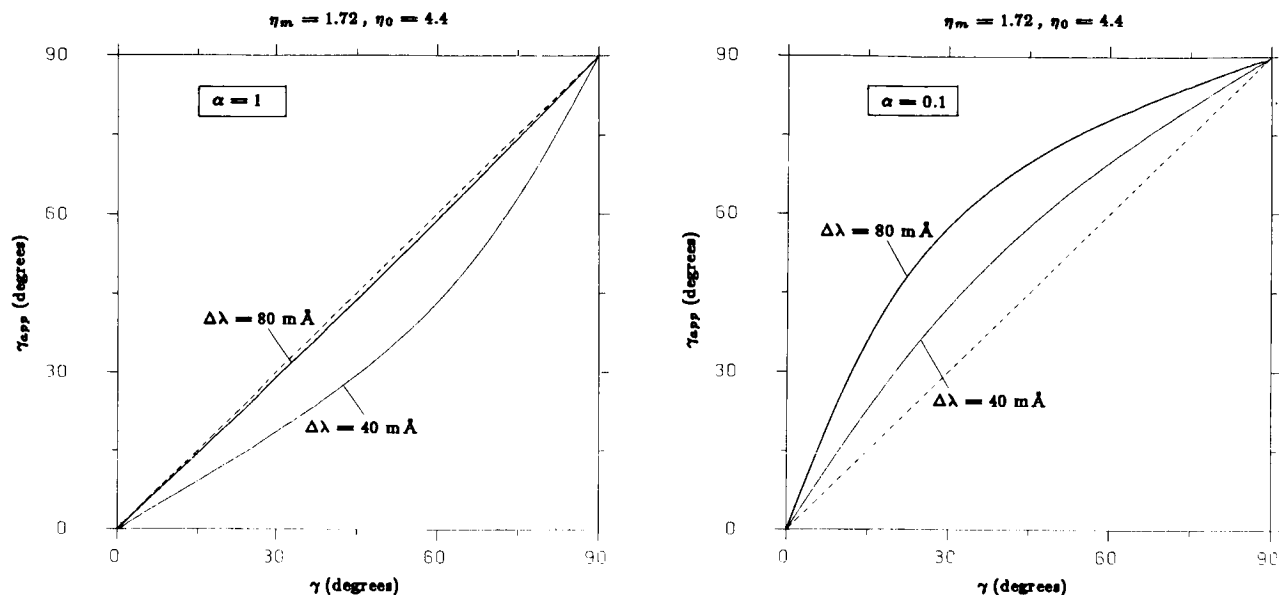


Fig. 6. Apparent angle of inclination γ_{app} between the field and the line of sight vs. the true angle γ , for area factor $\alpha = 1$ (left diagram) and $\alpha = 0.1$ (right diagram), including line weakenings in both cases.

field outside sunspots to look much more horizontal than it really is. Still another demonstration of the effect is given in Fig. 8, giving the apparent inclination γ_{app} as a function of the average longitudinal field $\alpha B_{\parallel} = \alpha B \cos 45^\circ$, assuming that the true inclination is fixed at $\gamma = 45^\circ$. Instead of running near the horizontal dashed line, all the curves start off at $\gamma_{app} = 90^\circ$ and then converge more or less well around the 45° level. The diagrams give the results for a filter centered at $\Delta\lambda = 80 \text{ mÅ}$, but the calculations

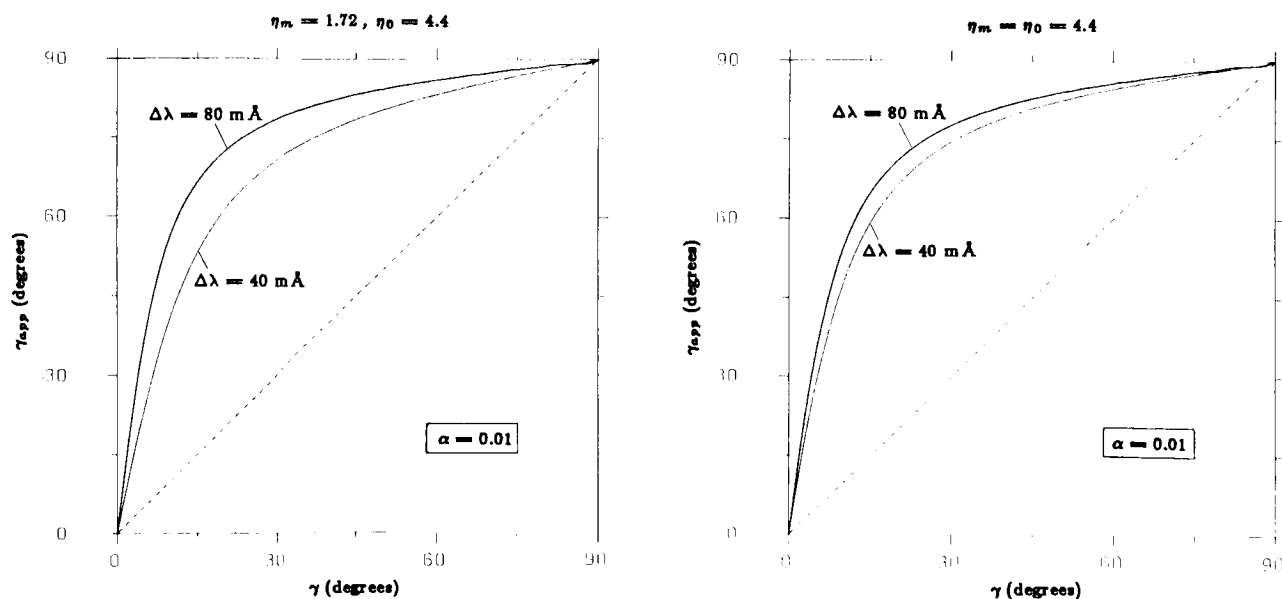


Fig. 7. γ_{app} vs. γ for $\alpha = 0.01$. Left (right) diagram: With (without) line weakenings.

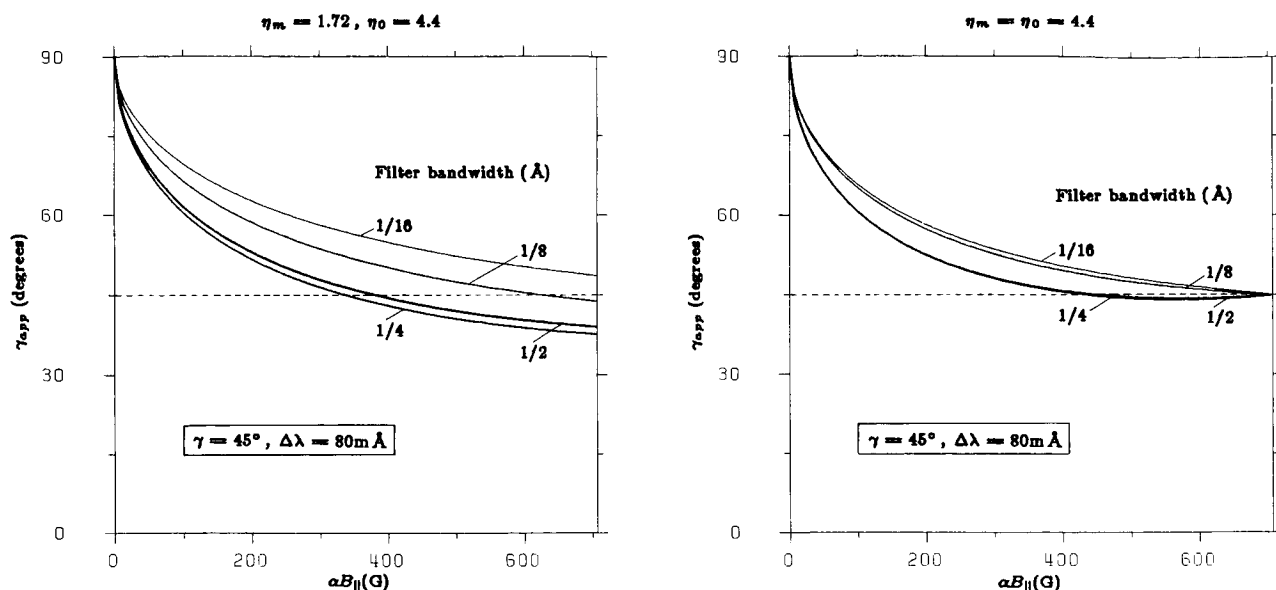


Fig. 8. Apparent field inclination γ_{app} vs. true average longitudinal field $\alpha B_{||}$, assuming a fixed true inclination $\gamma = 45^\circ$ and a filter position at $\Delta\lambda = 80 \text{ m}\text{\AA}$. The influence of the filter bandwidth is indicated by plotting curves for 1/16, 1/8, 1/4, and 1/2 Å filters. Left (right) diagram: With (without) line weakenings.

have been done for a series of filter bandwidths, 1/16, 1/8, 1/4, and 1/2 Å, to test the influence of the bandwidth. The temperature line weakening causes a spread in the curves for the various filters, but the main behaviour of the curves is dictated by α .

All the above calculations have been made including magneto-optical effects, but the results would not be changed if these effects are omitted.

4. Diagnostics with Spectrally Resolved Data

With the advent of the Fourier Transform Spectrometer (FTS) of the NSO McMath telescope and its conversion into a polarimeter, solar physics has got an instrument of enormous potential for fluxtube diagnostics. So far only FTS data for the I and V Stokes parameters have been analysed, but we have recently (May 3 – 7, 1984) made extensive FTS recordings of the three Stokes parameters, I , Q , and V , including their center-to-limb variations, both in the visible and in the infrared.

Fig. 9 shows a small portion of an FTS recording in a plage near disk center, made in April 1979 (Stenflo et al., 1984). It includes the line pair 1 Fe I 5247.06 and 5250.22 Å, which has been used in the past to diagnose the kG field strengths in the quiet network (Stenflo, 1973) and in plages (Frazier and Stenflo, 1978). If the fields were intrinsically weak, the ratio between the polarization amplitudes of these two lines should be 2 : 3, the ratio between their Landé factors, according to (2.6) and (2.4), but a glance at Fig. 9 shows that the ratio is much closer to 1:1. Using this deviation from the Landé-factor ratio, the intrinsic field strengths can be derived.

Fig. 10, from Stenflo and Harvey (1984), shows the observed relation between the polarization amplitudes in these two lines. The filled squares represent observations with the vertical grating spectrometer of the McMath telescope, the crosses FTS observations. The systematic difference between

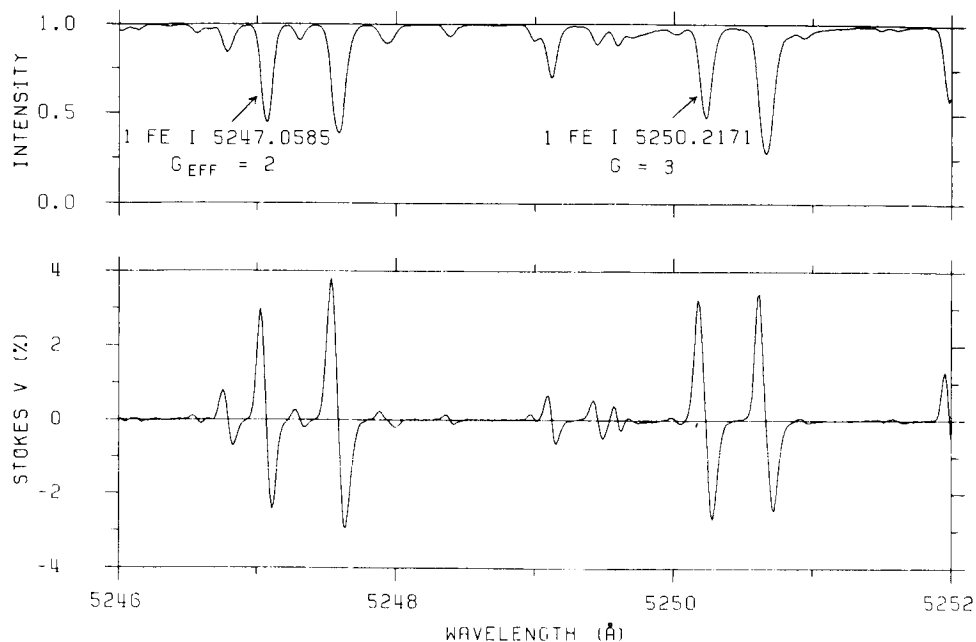


Fig. 9. Stokes I and V spectra around the 5247 and 5250 Å lines, recorded in a strong plage near disk center with the NSO McMath FTS polarimeter.

the FTS and grating spectrometer data is probably due to the limited spectral resolution (20 mÅ) and spectrometer straylight ($\sim 8\%$) in the grating spectrometer data, whereas these instrumental effects are absent in the FTS data.

The amount of polarization is proportional to the area factor α . Converting our FTS line ratios to

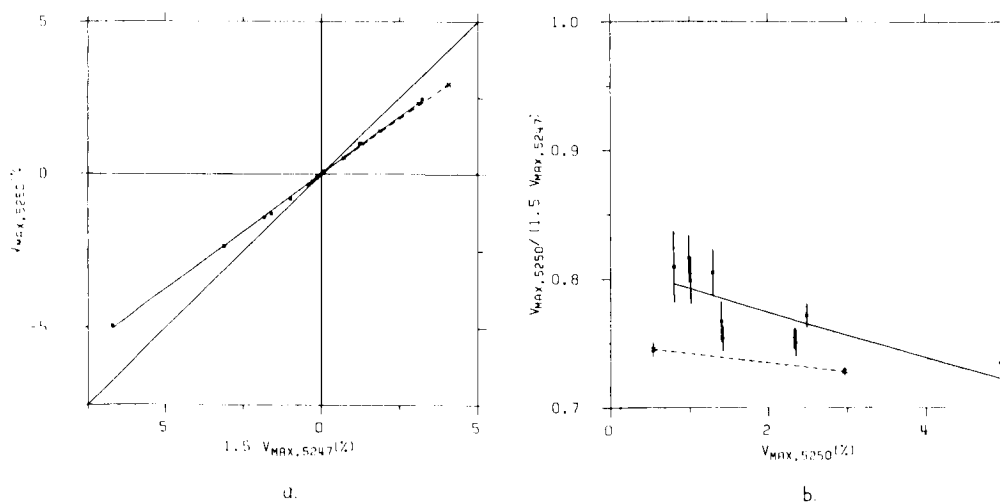


Fig. 10. (a) Polarization amplitude in Fe I $\lambda 5250.22$ Å vs. 1.5 times the polarization amplitude in Fe I $\lambda 5247.06$ Å. If the fields were intrinsically weak, the points would fall along the 45° line. (b) Ratio of the polarization amplitudes in the two lines vs. polarization amplitude in the 5250 Å line. Filled squares and solid curve (cubic spline fit): Spectrometer data. Crosses and dashed line: FTS data.

field strengths using our two-component model and Milne-Eddington line parameters, we find that the field strength increases from 1.0 to 1.1 kG when the area factor α increases by a factor of 6, thus a very slight dependence on α . This almost vanishing dependence is confirmed when we use a statistical approach with 400 unblended Fe I lines instead of just two (see below). The limited spectral resolution of the grating spectrometer data makes the field look apparently weaker than it really is.

A striking feature of Fig. 10 is the small spread of the points around the straight line in Fig. 10a, hardly exceeding the instrumental scatter. The error bars in Fig. 10b represent the standard deviations due to random instrumental noise, but do not include systematic effects like instrumental drifts or variable seeing. The well defined line ratio with no population of points in the upper part of Fig. 10b is empirical evidence for the applicability of the two-component model that we have used.

Another important feature of our FTS recordings can be seen by inspecting Fig. 9: the Stokes V line asymmetries. The amplitude a_b of the blue-wing polarization peak is always larger than the amplitude a_r of the red-wing polarization peak. However, also the area A_b of the blue V wing differs from that of the red V wing (A_r). In Fig. 11, from Solanki and Stenflo (1984), we have plotted the relative amplitude asymmetry, defined as $(a_b - a_r)/(a_b + a_r)$, and the relative area asymmetry, defined as $(A_b - A_r)/(A_b + A_r)$, vs. line strength S_I (Fraunhofer), defined as the area of the intensity profile below the half-level chord. The diagram is based on an analysis of a sample of 400 unblended Fe I lines in the visible part of the spectrum. The shaded areas are bounded by the envelopes of the smoothed asymmetry curves of all our observed magnetic regions. For the amplitude asymmetry, the network curves lie above the dashed line, the plage curves below it. For the area asymmetry no such distinction can be made.

Such asymmetries can only be explained if there are mass motions inside the magnetic fluxtubes. To produce the area asymmetry, a correlation between the magnetic-field and velocity gradients along the line of sight is required. One would then be tempted to try to explain the asymmetries in terms of fluxtube downdrafts in a diverging magnetic-field geometry.

However, any such model assuming some kind of quasi-stationary flow runs into serious difficulties,

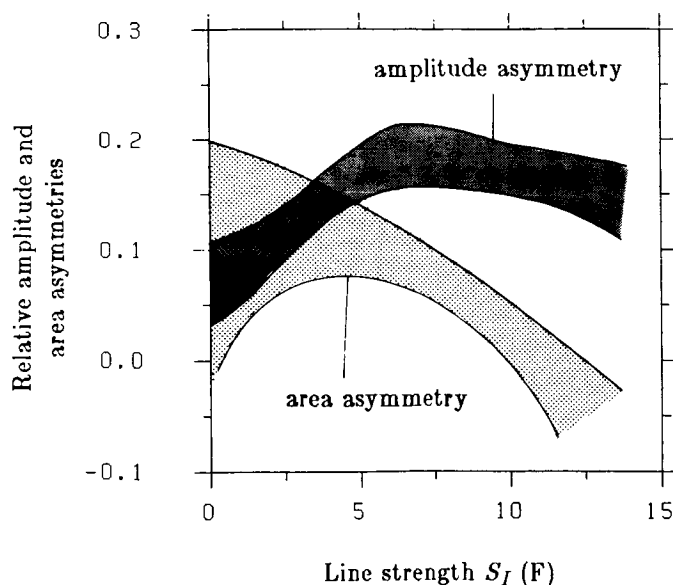


Fig. 11. Relative amplitude and area asymmetries in the Stokes V line profiles, plotted vs. line strength S_I . For details, see text.

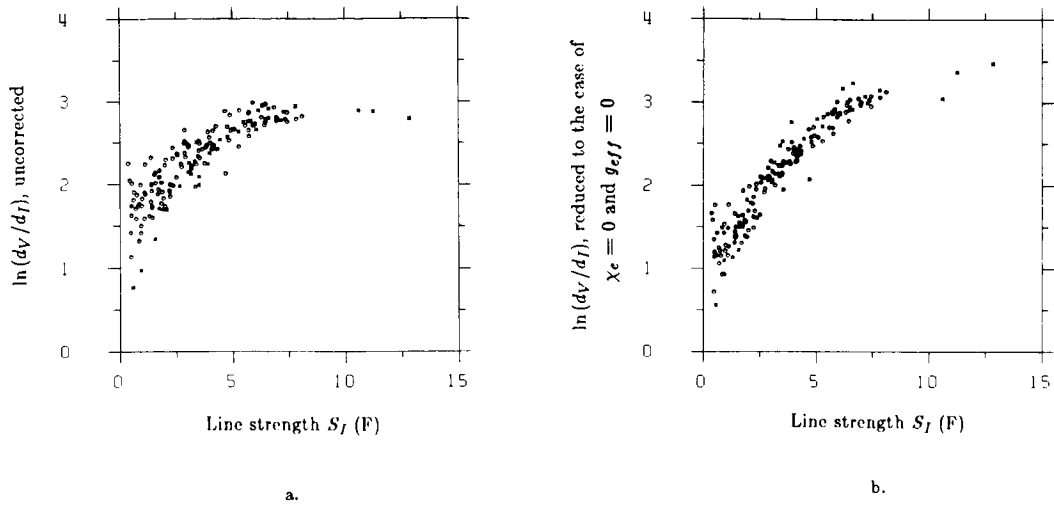


Fig. 12. $\ln(d_V/d_I)$ vs. S_I for our sample of Fe I lines (for definitions see text). The data were obtained with the FTS polarimeter in an enhanced network region near disk center. (a) Raw data. (b) the data after the application of eq. (4.1) has reduced them to the case of zero excitation potential and Landé factor.

because it produces a net Doppler shift of the zero crossing of the Stokes V profile with respect to the position of the corresponding Stokes I profile, and such a shift is not observed in our FTS spectra. We are therefore inclined to believe that no model with stationary flows can reproduce the observations, but that the flow is basically non-stationary in nature. The first type of non-stationary flow that comes to mind is fluxtube oscillations. Numerical calculations of fluxtube collapse by Hasan (1984) show how an end state develops that is not stationary but is oscillating.

As the weak-field approximation (2.4) is a good assumption for the majority of spectral lines except those with large Landé factor and narrow line width (like the 5250 Å line), we can retrieve the line profile inside the fluxtube by integrating V in (2.6) over λ . The line depth d_V that we then derive scales with the unknown area factor α , but this scale factor is the same for all lines in the FTS recording. In the following diagrams we can regard the absolute scale for d_V as arbitrary.

While the source of d_V is the magnetic fraction α of the aperture, the relative line depth d_I of the intensity profile gets its main contribution from the non-magnetic fraction $1 - \alpha$ of the atmosphere if α is small. In Fig. 12a, from Solanki and Stenflo (1984), we have plotted $\ln(d_V/d_I)$ vs. line strength S_I for our sample of Fe I lines, using an FTS recording in an enhanced network near disk center. Spectral lines with excitation potential $\chi_e < 3\text{eV}$ are represented by a star, those with $\chi_e \geq 3\text{eV}$ by a circle.

Using a regression equation of the form

$$\ln(d_V/d_I) = x_1 + x_2 S_I + x_3 S_I^2 + x_4 \chi_e + x_5 S_I \chi_e + x_6 g_{eff}^2 \lambda^2 / v_0^2, \quad (4.1)$$

one can remove the effect of χ_e and effective Landé factor g_{eff} on $\ln(d_V/d_I)$, to obtain the plot in Fig. 12b, in which the scatter of the points has been substantially reduced. (v_0 in (4.1) represents an empirical Doppler width of the lines.)

The relative variation of $\ln(d_V/d_I)$ with line strength is caused by the height variation of the temperature enhancement ΔT inside the fluxtube with respect to the surroundings. In Fig. 13, from Solanki and Stenflo (1984), we compare the empirical smoothed averages for plages and network (thick lines) with computer calculations for a fluxtube model with ΔT being 500 K at the top of the photosphere, while at

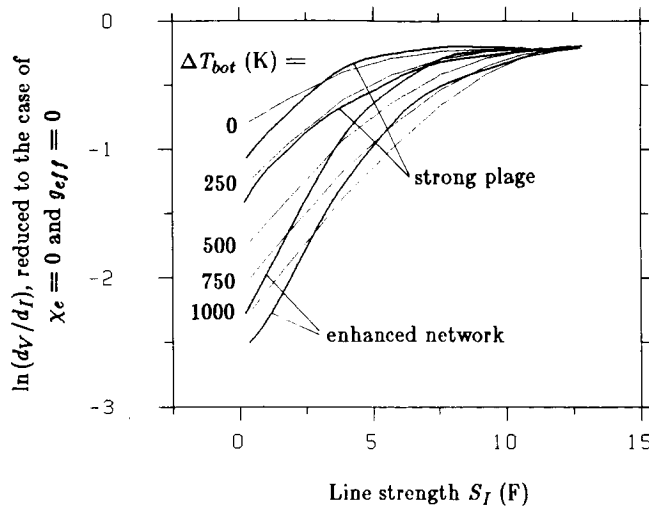


Fig. 13. $\ln(d_V/d_I)$ vs. S_I . The thick curves represent smoothed averages (cubic splines) of the data, reduced to the case of zero excitation potential and Landé factor. The thin lines are theoretical curves calculated using models with various temperature enhancements at the bottom of the photosphere (keeping the enhancement at the top of the photosphere fixed at $\Delta T = 500$ K). The empirical curves have been shifted in the vertical direction such that they coincide for the strongest spectral lines. (The shift corresponds to adjusting the value of the area factor α .)

the bottom of the photosphere $\Delta T = 0, 250, 500, 750$, and 1000 K, respectively. This comparison indicates that the network fluxtubes are substantially hotter than the plage fluxtubes in the lower photosphere.

Thus while a clear difference in temperature structure between plages and network is found, a statistical analysis of the line parameters plotted vs. $g_{eff}^2 \lambda^2 / v_0^2$ tells us that the intrinsic field strength is practically the same, which is also consistent with our previous discussion of Fig. 10.

5. Conclusions

We have shown how the vector magnetic fields derived from observations with a filter magnetograph will be severely distorted if the spatially unresolved magnetic structure is not properly accounted for. Thus the apparent vector field will appear much more horizontal than it really is, but this distortion is strongly dependent on the area factor and the temperature line weakenings. As the available fluxtube models are not sufficiently well determined, it is not possible to correct the filter magnetograph observations for these effects in a reliable way, although a crude correction is of course much better than no correction at all.

The solution to this diagnostic problem is to observe simultaneously in suitable combinations of spectral lines, and/or use Stokes line profiles recorded with very high spectral resolution. We have indicated the diagnostic power of using a Fourier transform spectrometer for polarimetry, and illustrated some results from I and V spectra.

The line asymmetries caused by mass motions inside the fluxtubes adds an extra complication to the diagnostic problem, in particular as there are indications that the motions are non-stationary in nature. The temperature structure appears to be a function of fluxtube diameter, as a clear difference between plage and network fluxtubes has been revealed.

The divergence of the magnetic field with height plays an essential role in the explanation of the Stokes V asymmetries (in combination with the mass motions). A self-consistent treatment of the subarcsec

field geometry may be required to allow an accurate derivation of the spatially averaged vector magnetic field from spectrally resolved data.

We are now extending our work with the McMath FTS polarimeter and have made simultaneous recordings of Stokes I , Q , and V spectra in various magnetic regions at different center-to-limb distances, with the aim of systematically building self-consistent fluxtube models, and to establish a diagnostic foundation for the derivation of vector magnetic fields from Stokes measurements.

References

- Baur, T.G., House, L.L., Hull, H.K.: 1980, *Solar Phys.* **65**, 111.
Baur, T.G., Elmore, D.E., Lee, R.H., Querfeld, C.W., Rogers, S.R.: 1981, *Solar Phys.* **70**, 395.
Chapman, G.A.: 1977, *Astrophys.J.Suppl.* **33**, 35.
Chapman, G.A.: 1979, *Astrophys.J.* **232**, 923.
Frazier, E.N., Stenflo, J.O.: 1978 *Astron.Astrophys.* **70**, 789.
Hagyard, M.J., Cumings, N.P., West, E.A., Smith, J.E.: 1982 *Solar Phys.* **80**, 33.
Harvey, J.W.: 1977, in E.A. Müller (ed.), *Highlights of Astronomy* **4**, 223.
Hasan, S.: 1984, *Astrophys.J.*, submitted.
House, L.L., Steinitz, R.: 1975, *Astrophys.J.* **195**, 235.
Kotov, V.A.: 1972, *Izv.Krymsk.Astrofiz.Obs.* **46**, 115.
Rachkovsky, D.N.: 1962a, *Izv.Krymsk.Astrofiz.Obs.* **27**, 148.
Rachkovsky, D.N.: 1962b, *Izv.Krymsk.Astrofiz.Obs.* **28**, 259.
Severny, A.B.: 1964, *Izv.Krymsk.Astrofiz.Obs.* **31**, 126.
Severny, A.B.: 1965, *Izv.Krymsk.Astrofiz.Obs.* **33**, 3.
Solanki, S.K., Stenflo, J.O.: 1984, *Astron.Astrophys.*, submitted.
Stenflo, J.O.: 1971, in R. Howard (ed.), 'Solar Magnetic Fields', *IAU Symp.* **43**, 101.
Stenflo, J.O.: 1973, *Solar Phys.* **32**, 41.
Stenflo, J.O.: 1975, *Solar Phys.* **42**, 79.
Stenflo, J.O.: 1976, in V. Bumba, J. Kleczek (eds.), 'Basic Mechanisms of Solar Activity', *IAU Symp.* **71**, 69.
Stenflo, J.O., Harvey, J.W.: 1984, *Solar Phys.*, submitted.
Stenflo, J.O., Harvey, J.W., Brault, J.W., Solanki, S.K.: 1984, *Astron.Astrophys.* **131**, 333.
Stenholm, L.G., Stenflo, J.O.: 1978, *Astron.Astrophys.* **67**, 33.
Stepanov, V.E.: 1958a, *Izv.Krymsk.Astrofiz.Obs.* **18**, 136.
Stepanov, V.E.: 1958b, *Izv.Krymsk.Astrofiz.Obs.* **19**, 20.
Stepanov, V.E., Severny, A.B.: 1962, *Izv.Krymsk.Astrofiz.Obs.* **28**, 166.
Unno, W.: 1956, *Publ.Astron.Soc.Japan* **8**, 108.

**Generation and Transfer of Polarized Radiation in the Solar Atmosphere:
Physical Mechanisms and Magnetic-Field Diagnostic**

E. Landi Degl'Innocenti

Istituto di Astronomia, Università di Firenze

Abstract

The main physical mechanisms responsible for the generation and transfer of polarized radiation in the solar atmosphere can be classified in a suitable bidimensional diagram with an indicator of the magnetic field strength on its vertical axis and an indicator of the radiation field anisotropy on its horizontal axis.

The various polarimetric observations performed on solar spectral lines are interpreted with different theoretical schemes according to their classification in the diagram and to the optical depths involved.

These theoretical schemes, and the associated diagnostic tools for inferring the magnetic field vector from observations are reviewed.

In particular, the role of magneto-optical effects in determining the direction of the observed linear polarization in active regions is discussed in some detail.

1. Physical Mechanisms

Measurements of solar magnetic fields are generally deduced from direct observations of polarization in spectral lines. For the correct diagnostic of solar magnetic fields it is then fundamental to have a detailed understanding of the physical mechanisms that are able to generate polarized radiation and those that are effective in modifying its properties in the transfer through the solar atmosphere.

Polarization can be originated in solar spectral lines by means of two different physical mechanisms: a) the Zeeman effect and b) the presence of atomic polarization.

The Zeeman effect is the familiar phenomenon which is at the basis of the measurement of solar magnetic fields by means of standard magnetographs (longitudinal or vector magnetographs). The important parameter which determines the amplitude of the observed polarization due to the Zeeman effect is $v_H = g \nu_L / \Delta \nu_D$ where g is the Landé factor, ν_L is the Larmor frequency and $\Delta \nu_D$ the Doppler broadening of the spectral line; for small values of the magnetic field the circular polarization is proportional to v_H , while the linear polarization is proportional to v_H^2 , but these simple relationships rapidly break down as v_H increases and approaches values of the order of unity.

Atomic polarization, on the contrary, is a phenomenon that is more familiar to spectroscopists working in optical pumping experiments. It consists in the fact that, when a degenerate atomic level is excited by non-thermal mechanisms, its sublevels are in general, not evenly populated and may show well defined phase relationships (often called interferences or coherences) among them. From the solar physicist's point of view, atomic polarization is a typical non-LTE phenomenon; it can be induced in the solar atmosphere by the radiation field anisotropy or, alternatively by collisions with oriented beams of fast particles. This last case however, that is often referred to as impact polarization, will not be considered in the present paper.

Atomic polarization can be described in several equivalent ways; the most familiar is the one that employs the density matrix elements between magnetic sublevels of the same energy level; namely

$$\rho(M, M') = \langle \alpha J M | \rho | \alpha J M' \rangle$$

where ρ is the quantum-mechanical density matrix operator and $|\alpha J M\rangle$ is the eigenvector of the atomic Hamiltonian. The diagonal elements, $\rho(M, M)$, represent the populations of the single magnetic sublevels while the off-diagonal elements represent the so-called coherences between different sublev-

els. By means of suitable linear combinations of the density matrix elements, two important parameters can be defined; for the case of a level with $J=1$ we have, for example:

$$p_1 = [\rho(1, 1) - \rho(-1, -1)] / [\rho(1, 1) + \rho(-1, -1)]$$

$$p_2 = [\frac{1}{2}(\rho(1, 1) + \rho(-1, -1) - \rho(0, 0))] / [\rho(1, 1) + \rho(0, 0) + \rho(-1, -1)]$$

These two parameters are called, respectively, the orientation and alignment parameters; roughly speaking the radiation emitted in a line whose upper level is characterized by certain values p_1 and p_2 will show an amount of fractional polarization given by:

$$P_{\text{or}} \approx p_1 \quad P_{\text{ln}} \approx p_2$$

While there is no known physical mechanism able to induce atomic orientation in the atoms of the solar atmosphere, atomic alignment can be easily produced by the anisotropy of the radiation field. This anisotropy can in turn be due to geometrical effects or to limb-darkening; a particular example is given by the higher layers of the solar atmosphere (chromosphere and corona), whose atoms are illuminated by the photospheric radiation field and show the well known phenomenon of resonance polarization (limb polarization).

To give an idea of the order of magnitudes involved, the resonance polarization observed in a spectral line can be written in the form:

$$P_{\text{ln}} \approx \frac{R_a}{R_i + D}$$

where R_i and R_a are the isotropic and anisotropic radiative rates, respectively, and where D is the depolarizing collisional rate. While R_i is simply given by the expression:

$$R_i = B \int I(\vec{\Omega}) d\vec{\Omega}$$

where B is the Einstein coefficient for radiative excitation from lower to upper

level, R_a has in general a more involved expression and is often written in the form:

$$R_a = W_a R_i$$

where W_a is the so-called anisotropic dilution factor. For a completely isotropic radiation field W_a is zero, while, for a unidirectional radiation beam, W_a is equal to unity. In intermediate situations the expression for W_a is more involved; supposing that the atom is illuminated by a radiation field which fills uniformly a cone having semiaperture θ_0 , the expression for W_a is the following:

$$W_a = \frac{1}{2} \cos \theta_0 (1 + \cos \theta_0)$$

When a magnetic field is present in the region where resonance radiation is scattered, there is a different effect that is able to affect the signature of the emitted polarization. This is the Hanle effect which consists in the fact that a given magnetic field can reduce or, in some cases, even completely destroy the phase relationships between degenerate or quasi-degenerate sublevels. The basic physical mechanism responsible for the Hanle effect can be easily understood by means of a semiclassical analogy which consists in describing the atoms as three-dimensional oscillators for deducing the results to be expected in typical scattering experiments.

With reference to figure 1, suppose that the scattering atom is illuminated by an unpolarized pencil of radiation propagating along the y axis. If the atom is described as a collection of three linear oscillators directed along the x, y and z axis, the unpolarized radiation beam will excite the x and z oscillators without any definite phase relationship between them. Observing along the x direction we will then obtain a light beam that is totally polarized along the z-axis as the x-oscillator, being viewed head-on, will not contribute to the scattered radiation.

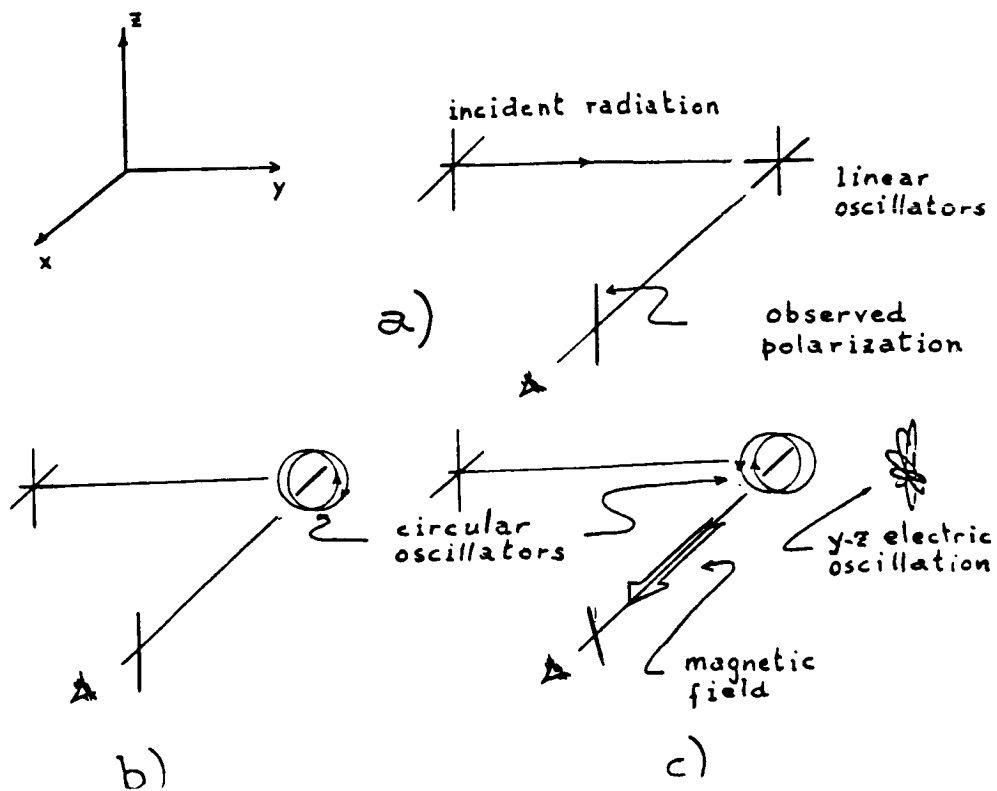


Figure 1

The description of the atom in terms of three linear oscillators as given in Fig. 1a is, however, an arbitrary one. Alternatively, one may choose to describe the atom in terms of one linear oscillator directed along the x-axis and two circular oscillators, one right-handed and the other left-handed, both laying in the y-z plane (see Fig 1b). In this description, the incident radiation beam will

excite again the x-oscillator and will also excite the two circular oscillators, though with a well defined phase relationship between them; the right-handed and left-handed circular oscillators will move in such a way that the resulting oscillation will take place along the z-axis. The polarization observed along the x direction is obviously the same as derived previously but we see that with the description of the atom given in Fig. 1b, we are obliged to take properly into account the phase relationships between different oscillators to have a suitable description of the atomic excitation. These phase relationships between classical oscillators represent, in an intuitive way, the off-diagonal elements of the atomic-density matrix that are needed in the quantum-mechanical description of the atom.

If we now introduce a magnetic field along the x-axis, the total arbitrariness, that was previously present for choosing the description of the atomic excitation in terms of classical oscillators, is lost. The most convenient choice is the one which utilizes a linear oscillator along the direction of the magnetic field and two circular oscillators in the perpendicular plane or, in other words, the same as in Figure 1b. The difference with the non-magnetic case is contained in the fact that, due to Larmor precession, the two circular oscillators have now different frequencies and, as a result, a well defined phase relationship cannot be kept indefinitely between them (Hanle effect). As the difference in frequency is of the order of ν_L and the lifetime of the oscillators is $1/A$ (where A is the Einstein coefficient for spontaneous deexcitation) the critical parameter for the Hanle effect is the ratio ν_L/A . If $\nu_L/A \ll 1$ the phase relationship between the two circular oscillators can be kept indefinitely and the polarization observed in the scattering experiment is the same as in the non-magnetic case; vice versa if $\nu_L/A \gg 1$ the phase relationship is lost so that the two circular oscillators behave independently and the polarization is destroyed. Finally, in the intermediate situation where $\nu_L/A \approx 1$ the oscillation will start

along the z-direction but will then describe a complex path in the form of a rosette as depicted in Fig. 1c; as a result, the linear polarization will be partly reduced and tilted with respect to the non-magnetic situation.

Summarizing the discussion presented above, the various physical mechanisms that are responsible for the generation of polarization in solar spectral lines can be classified in a suitable bidimensional diagram having an indicator of the magnetic field intensity on its vertical axis and an indicator of the radiation field anisotropy on its horizontal axis. This diagram has been presented and discussed elsewhere (Landi Degl'Innocenti, 1983b) and the underlying theory, which can be applied to the different physical regimes, has been developed in full generality by Landi Degl'Innocenti (1983a). The main advantage of this theoretical scheme is its unified approach which allows the formalism to be applied to the different physical regimes that are met in the solar atmosphere. In a recent paper, for instance, the general formalism has been applied to describe the physics of resonance polarization in the zero-magnetic field, collisionless regime (Landi Degl'Innocenti, 1984).

Once polarized radiation has originated in a particular region of the solar atmosphere, its properties can be deeply modified through transfer effects in the atmosphere itself. A very general transfer equation for polarized radiation in the presence of both Zeeman splitting and atomic polarization has been given by Landi Degl'Innocenti (1983a). In the case where atomic polarization can be neglected the radiative transfer equations reduce to the well known expression often referred to as Unno-Rachkowsky equations (Unno, 1958; Rachkowsky, 1962):

$$\frac{d}{ds} \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = - \begin{pmatrix} \eta_I & \eta_Q & \eta_U & \eta_V \\ \eta_Q & \eta_I & \rho_V & -\rho_U \\ \eta_U & -\rho_V & \eta_I & \rho_Q \\ \eta_V & \rho_U & -\rho_Q & \eta_I \end{pmatrix} \begin{pmatrix} I-S \\ Q \\ U \\ V \end{pmatrix} \quad (1)$$

where the meanings of the various symbols and the conventions for the signs of the Stokes parameters can be found in full detail in Landi Degl'Innocenti (1976) or in Landolfi and Landi Degl'Innocenti (1982).

The absorption matrix in equation (1) contains the terms ρ_Q, ρ_U, ρ_V that are responsible for magneto-optical effects. During the last two decades the importance of magneto-optical effects in the physics of radiative transfer for polarized radiation has been widely debated. Summarizing some theoretical and observational results that have been obtained more recently, we can derive the following qualitative conclusions: i) magneto-optical effects can be neglected for $\Psi = 0^\circ$ or $\Psi = 90^\circ$, where Ψ is the inclination angle of the vector magnetic field with respect to the line of sight; ii) magneto-optical effects play a negligible role in the limiting cases of very weak or very strong magnetic field ($v_H \ll 1$ and $v_H \gg 1$, respectively); iii) for intermediate values of the magnetic field ($v_H \sim 2$) and large inclination angles ($40^\circ \lesssim \Psi \lesssim 80^\circ$) magneto-optical effects induce a reversal in the V-Stokes profile around line center; iv) magneto-optical effects have considerable influence on the signature of linear polarization profiles (especially near line-center) and affect dramatically the value of the integrated linear polarization over the whole spectral line. Conclusions i) and ii) can be easily proven by means of the geometrical interpretation of the equations of radiative transfer given by Landi Degl'Innocenti and Landi Degl'Innocenti (1981); in the cases $\Psi = 0^\circ$ or $\Psi = 90^\circ$ the formal vectors $\vec{\eta} = (\eta_Q, \eta_U, \eta_V)$ and $\vec{\rho} = (\rho_Q, \rho_U, \rho_V)$ are parallel and the influence of magneto-optical effects is rigorously zero; the same parallelism is approximately fulfilled also for $v_H \gg 1$ and $v_H \ll 1$ (Landi Degl'Innocenti and Landi Degl'Innocenti, 1973). Conclusion iii) has been suggested by Wittmann (1971) and has been subsequently confirmed by various investigations; finally, conclusion iv) has been directly proven by comparing observations in sunspots with theoretical expectations from suit-

able models. (Calamai et al., 1975; Landi Degl'Innocenti, 1979; West and Hagyard, 1983; Kawakami, 1983).

2. Magnetic Field Diagnostic

Although the basic physical mechanisms that are involved in the generation and transfer of polarized radiation in the solar atmosphere seem nowadays to be fairly well understood, the inverse problem of deducing the value of the magnetic field vector from polarimetric observations is still far from being adequately solved. There is no doubt that this is due to the involved structure of the solar atmosphere, especially in those regions that are typically investigated through polarimetric observations. Most of the diagnostic tools that have been developed for deducing the value of the magnetic field vector from polarimetric observations are indeed based on the rather crude assumption that schematizes the line-forming region as a unidimensional atmosphere embedded in a magnetic field that is supposed constant with depth and uniform over the observed area. This assumption is clearly a non-realistic one for the interpretation of polarimetric observations in plages or in the photospheric network as it is nowadays well established that in these regions the magnetic field is concentrated in structures having dimensions well below the resolving power of ground based instruments. The diagnostic problems associated with these observations are discussed in full detail by Stenflo (1984) and will not be addressed in this paper.

2a. Diagnostic of Magnetic Fields in Prominences

Observations of linear and circular polarization in the HeI D_3 line in prominences have revealed to be particularly suitable for the measurement of vector magnetic fields through the Hanle effect. The theoretical interpretation has

been given, through similar but independent methods, by Bomnier (1980) and Landi Degl'Innocenti (1982). The diagnostic method consists in solving the statistical equilibrium equations for the density matrix of the triplet system of the HeI atom. The atom, located at a certain height h over the solar limb, is illuminated by the photospheric radiation field and its atomic polarization is strongly affected by the presence of a magnetic field. From the solution of the statistical equilibrium equations, the emissivity in the D_3 line is derived, for each of the Stokes parameters, as a function of the relevant parameters:

$$\begin{pmatrix} \varepsilon_I \\ \varepsilon_Q \\ \varepsilon_U \\ \varepsilon_V \end{pmatrix} = \Phi_\lambda(h, B, \theta, \phi)$$

where B is the amplitude of the magnetic field vector and where θ and ϕ are the angles that specify its direction with respect to the solar radius through the prominence.

In principle, if h is known from observations, the values of B , θ and ϕ can be recovered by direct comparison of the theoretical results with the observed polarization as done, for instance, by Athay et al. (1983). For this inversion to be possible it is necessary to know from observations the linear polarization in the two fine-structure components of the D_3 line. From broad-band D_3 observations, as the ones performed with the Pic-du-Midi polarimeter, it is only possible to deduce two parameters of the magnetic field vector, the third being supposed known a-priori. Examples of this inversion technique have been given by Leroy et al. (1983, 1984). In these papers, the amplitude B and azimuthal ϕ of the magnetic field vector are recovered supposing this vector perpendicular to the solar radius, an assumption that seems fairly well established at least for quiescent prominences (Leroy, 1979; Athay et al., 1983).

The theory of the Hanle effect, as developed for D_3 , can be extended to other spectral lines with the aim of obtaining a well-established set of theoretical expectations for observations that might be available in the near future. Bommier and Sahal-Brechot (1982) have investigated the linear polarization of $Ly\alpha$ in the corona, Sahal-Brechot et al. (1984) have addressed the problem of the polarization of $O\ VI\ \lambda\ 1032$ in the transition region and Landolfi and Landi Degl'Innocenti (1984) have derived the expected polarization of the $Na\ I\ D_2$ line in prominences as a function of the magnetic field.

The inversion technique for deducing the magnetic field vector through the Hanle effect suffers from the main limitation that, for optically thin lines, there is an intrinsic symmetry which inhibits from discriminating between a value of \vec{B} and its symmetric with respect to the line of sight. Recent calculations for the optically thick $H\alpha$ line in prominences (Bommier et al., 1984) have however shown that this symmetry is broken as a result of multiple scattering effects, so that simultaneous observations in D_3 and $H\alpha$ should be able of yielding a unique determination of the magnetic field vector in prominences through the Hanle effect.

2b. Diagnostic of Magnetic Fields From Disk Observations

For disk observations it is necessary to make a distinction among three different types of observing techniques namely:

1) broad-band or white light observations, 2) filter-type or magnetograph-type observations, 3) narrow-band or profile-type observations.

These three classes of observations are characterized by decreasing values of spectral resolution $\Delta\lambda$; schematically, in broad-band observations $\Delta\lambda$ is larger or of the order of 50 to 100 Å so that many spectral lines are present in the

observed wavelength range; in filter-type or magnetograph-type observations $\Delta\lambda \simeq 100m\text{\AA}$, and, generally, only few wavelength points are sampled across the profile of a magnetic sensitive line; finally, in profile-type observations $\Delta\lambda \simeq 30m\text{\AA}$ and a complete coverage of the line profile is performed.

2b.1. White Light Observations

After the discovery of Dollfus (1958), broad-band linear polarization in sunspots was extensively observed by Leroy (1962) and later confirmed by Illing et al. (1974a,b; 1975). These last authors also discovered the presence of broad-band circular polarization in sunspots, a rather unexpected phenomenon that was more recently investigated in greater detail by Kemp and Henson (1983).

The interpretation of the observed linear polarization was given by Leroy (1962) in terms of a simplified theory which invoked the mechanisms of magnetic intensification as the responsible agent for the appearance of this phenomenon. Leroy's theory was subsequently generalized to take properly into account the influence of magneto-optical effects that play, in this case, a very important role as it has been discussed in Section 1; the theoretical results concerning the magnetic intensification mechanism as modified by magneto-optical effects have been recently presented in the forms of tables and graphs by Landi Degl'Innocenti and Calamai (1982) and Calamai and Landi Degl'Innocenti (1983).

Although the basic physical processes involved are fairly well understood, the diagnostic content of broad-band linear polarization observations is rather small and is strictly limited to sunspots. The amplitude of the magnetic field vector remains highly undetermined while its direction can be roughly recovered if the polarity of the spot is known by means of complementary

observations.

Broad-band circular polarization observations seem to be more promising as a diagnostic tool for inferring velocity gradients in sunspots. Velocity gradients (in combination or not with magnetic field gradients) may indeed be responsible for the appearance of net circular polarization in sunspots, the requirement of having both kind of gradients being necessary only for normal triplets in a longitudinal magnetic field. Despite some qualitative results obtained by Auer and Heasley (1978) and by Landman and Finn (1979) the interpretative problem of broad-band circular polarization in sunspots is still an open one and needs further investigation before it may acquire the role of a quantitative diagnostic tool.

2b.2. Magnetograph-Type Observations

As the prototype of a magnetograph-type observation we consider the idealized case of an instrument giving, for the measurement of the Stokes parameters, the four quantities:

$$\tilde{S}_i(\lambda_0) = \int S_i(\lambda) f(\lambda - \lambda_0) d\lambda$$

where $f(\lambda - \lambda_0)$ is the filter profile centered at the wavelength λ_0 and having a typical width of the order of 100 m Å.

From the observed values $S_i(\lambda_0)$ (in some cases even repeated for some different wavelengths λ_0) one needs to recover the magnetic field vector, or, in other words, one needs to have a suitable *calibration* of the magnetograph.

For longitudinal magnetographs the component of the magnetic field along the line of sight, $B_{||}$, is usually recovered according to the classical formula:

$$V = k B_{||} g \frac{dI}{d\lambda} \quad (2)$$

where g is the Landé factor of the line and k is a constant depending on the instrumental apparatus. This formula has very serious limitations that are due both to the fact that the linear relationship between V and $dI/d\lambda$ is rigorously valid only in the limiting case of very weak magnetic fields and, moreover, to the fact that for unresolved observations, the observed V profile is representative only of the magnetic regions where the thermodynamical structure of the atmosphere is generally different from the non-magnetic regions. Although the longitudinal magnetograph still remains a very useful tool, in as much as it makes possible the observation of the general configuration of the solar magnetic field at a given time, the two limitations described above prevent its possible use as a quantitative measuring device.

Some efforts have been made in recent times with the aim of obtaining a more precise calibration than the approximate one described by formula (2). As an example, Hagyard et al. (1982), and West and Hagyard (1983) have carefully described the calibration of the Marshall Space Flight Center vector polarimeter which is based on the Kjeldseth Moe (1968) analytical solutions to the radiative transfer equations. The real problem is, however, that there cannot be a unique calibration valid for all the physical situations that are met in the solar atmosphere; in particular, the role of velocity gradients (associated or not with magnetic field gradients) in the line forming regions and its influence on magnetographic vector field determinations has still to be clarified.

2b.3. Profile-Type Observations

These are the most complete observations, from the polarimetric point of view, that can be accomplished on magnetic sensitive lines. In particular, through such observations, it has been discovered that, in many cases, the Stokes profiles are largely asymmetric or may show very peculiar signatures. Some examples of these "anomalous" profiles have been shown by Mickey

(1984); their interpretation is still unknown but may well hide some very interesting phenomena in the physics of solar magnetic regions.

From Stokes profiles of magnetic lines (or at least from these profiles that do not show anomalous behavior) it is possible to deduce the magnetic field vector with a certain degree of confidence. To this aim several algorithms have been developed in recent times; these are the following: a) *Unno-fit technique*: the magnetic field vector is deduced by means of a non-linear least-squares fit to the analytical formulae of Unno (1956); the technique was firstly developed by Auer et al. (1978) and has been recently generalized to account for magneto-optical effects and damping effects (Landolfi and Landi Degl'Innocenti, 1982; Landolfi et al., 1984; Lites and Skumanich, 1984). b) *Lambdameter technique*: the longitudinal component of the magnetic field is deduced from the distance in wavelength between the centers of gravity of right and left circular polarization profiles (Semel, 1970; Rees and Semel, 1979); with this technique the transversal component of the magnetic field is not determined. c) *Fourier transform technique*: the wavelength profiles of the Stokes parameters are Fourier-transformed and the magnetic field components are measured through suitable properties of the Fourier-transforms of the Stokes profiles (Title and Tarbell, 1975; Tarbell and Title, 1976); this technique works properly only for magnetic fields $B \gtrsim 1000$ gauss. d) *Eigenvalue technique*: this technique is rather involved and is based on the fact that, under a certain number of assumptions that are the same as those of the Kjeldseth-Moe analytical solutions to the transfer equations, the magnetic field vector can be recovered by comparing certain combinations of the Stokes parameters at different wavelengths; details can be found in Makita (1979) and a direct application in Kawakami (1983); the eigenvalue technique suffers from the main disadvantage that magneto-optical effects cannot be accounted for in a consistent way. e) *Characteristic-features technique*: for a given line and a given atmospheric

model, some characteristic features of the Stokes profiles (like for instance the wavelength of the maximum of the V-profile, or the amplitude of the maximum itself, or the value of the Q-profile at line center) are connected, through simple relationships to the magnetic-field parameters. Once the theoretical calibration is performed, the magnetic field can be recovered through the observed values of the characteristic features (Ye Shi-hui and Jin Je-hai, 1983). This technique suffers from the main limitation of being somewhat model dependent and of yielding results that are very much affected by the presence of noise in the observations. f) *Seares formulae technique*: this technique is based on the direct application of the Seares formulae (Loftin, 1984); as such it has very deep limitations, being valid only for non saturated normal Zeeman triplets that lay on the linear part of the curve of growth.

All the algorithms listed above are well suited to obtain a measurement of the magnetic field vector (or, at least, one of its components) in as much as they give the correct answer for synthetic profiles obtained from undimensional model-atmospheres in a magnetic field that is supposed constant with depth and uniform over the observed area. When applying the algorithms to real solar observations we face however some severe interpretative problems on the meaning itself that can be attributed to the deduced values of the magnetic field vector; in fact, if we are in the presence of non-constant, non-uniform magnetic fields and if the atmosphere where the line is formed is a non-static, pluridimensional atmosphere, each of the methods will produce a different "ill-defined" mean as measurement of the magnetic field and the fact that different methods will produce different measurements is not surprising.

Any of the algorithms that deduce a *single value* for the magnetic field parameters from Stokes profiles is indeed too reductive. If the aim is simply the one of giving a sort of mean value for the magnetic field parameters in the observed region, then there is no need of profile-type observations as vector

magnetographs just do the same job with sufficient accuracy.

Yet, the amount of information that is contained in Stokes profiles is very large. To extract this information it is necessary to generalize the algorithms listed above so to deduce not only the "mean" magnetic field, but also some different parameter describing its depth or spatial variation together with velocity fields or atmospheric inhomogeneities.

Although this work has still to be started, the inversion methods based on the Unno-fit technique or the eigenvalue-technique seem to be the more promising for this kind of future developments.

References

- Athay, R.G., Querfeld, C.W., Smartt, R.N., Landi Degl'Innocenti, E., Bommier, V., 1983, *Solar Phys.*, **89**, 3.
- Auer, L.H., Heasley, J.N., 1978, *Astron. Astrophys.*, **64**, 67.
- Auer, L.H., Heasley, J.N., House, L.L., 1977, *Solar Phys.*, **55**, 47.
- Bommier, V., 1980, *Astron. Astrophys.*, **87**, 109.
- Bommier, V., Sahal-Brechot, S., 1982, *Solar Phys.*, **78**, 157.
- Bommier, V., Landi Degl'Innocenti, E., Leroy, J.L., Sahal-Brechot, S., 1984, this volume.
- Calamai, G., Landi Degl'Innocenti, E., 1983, *Astron. Astrophys. Suppl.*, **53**, 311.
- Calamai, G., Landi Degl'Innocenti, E., Landi Degl'Innocenti, M., 1975, *Astron. Astrophys.*, **45**, 297.
- Dollfus, A., 1958, *Comptes Rendue*, **246**, 3590.
- Hagyard, M.J., Cumings, N.P., West, E.A., Smith, J.E., 1982, *Solar Phys.*, **80**, 33.
- Illing, R.M.E., Landman, D.A., Mickey, D.L., 1974a, *Astron. Astrophys.*, **35**, 327.
- Illing, R.M.E., Landman, D.A., Mickey, D.L., 1974b, *Astron. Astrophys.*, **37**, 97.
- Illing, R.M.E., Landman, D.A., Mickey, D.L., 1975, *Astron. Astrophys.*, **41**, 183.
- Kawakami, H., 1983, *Publ. Astron. Soc. Japan*, **35**, 459.
- Kemp, J.C., Henson, G.D., 1983, *Astrophys. J. Lett.*, **266**, L69.
- Kjeldseth Moe, O., 1968, *Solar Phys.*, **4**, 267.
- Landi Degl'Innocenti, E., 1976, *Astron. Astrophys. Suppl.*, **25**, 379.
- Landi Degl'Innocenti, E., 1979, *Solar Phys.*, **63**, 237.
- Landi Degl'Innocenti, E., 1982, *Solar Phys.*, **79**, 291.
- Landi Degl'Innocenti, E., 1983a, *Solar Phys.*, **85**, 3.
- Landi Degl'Innocenti, E., 1983b, *Solar Phys.*, **85**, 33.

- Landi Degl'Innocenti, E., 1984, *Solar Phys.*, **91**, 1.
- Landi Degl'Innocenti, E., Calamai, G., 1982, *Astron., Astrophys. Suppl.*, **49**, 677.
- Landi Degl'Innocenti, E., Landi Degl'Innocenti, M., 1973, *Solar Phys.*, **31**, 299.
- Landi Degl'Innocenti, E., Landi Degl'Innocenti, M., 1981, *Il Nuovo Cimento*, **62B**, 1.
- Landolfi, M., Landi Degl'Innocenti, E., 1982, *Solar Phys.*, **78**, 355.
- Landolfi, M., Landi Degl'Innocenti, E., 1984, *Solar Phys.* (submitted).
- Landolfi, M., Landi Degl'Innocenti, E., Arena, P., 1984, *Solar Phys.*, in press.
- Landman, D.A., Finn, G.D., 1979, *Solar Phys.*, **63**, 221.
- Leroy, J.L., 1962, *Annal. Astrophys.*, **25**, 127.
- Leroy, J.L., 1979, in E. Jensen, P. Maltby and F.Q. Orrall (eds.), *"Physics of Solar Prominences"*, *IAU Colloq.* **44**, 87.
- Leroy, J.L., Bommier, V., Sahal-Brechot, S., 1983, *Solar Phys.*, **83**, 135.
- Leroy, J.L., Bommier, v., Sahal-Brechot, S., 1984, *Astron. Astrophys.*, **131**, 33.
- Lites, B.W., Skumanich, A., 1984, this volume.
- Loftin, T.A., 1984, this volume.
- Makita, M., 1979, *Publ. Astr. Soc. Japan*, **31**, 575.
- Mickey, D.L., 1984, this volume.
- Rachkovsky, D.N., 1962, *Izv. Krymsk. astrofiz. Obs.*, **27**, 148.
- Rees, D.E., Semel, M.D., 1979, *Astron. Astrophys.*, **74**, 1.
- Sahal-Brechot, S., Malinovksy, M., Bommier, V., 1984, *Astron. Astrophys.* (submitted).
- Semel, M.D., 1970, *Astron. Astrophys.*, **5**, 330.
- Stenflo, J.O., 1984, this volume.
- Tarbell, T.D., Title, A.M., 1976, *Solar Phys.*, **47**, 563.
- Title, A.M., Tarbell, T.D., 1975, *Solar Phys.*, **41**, 255.

Unno, W., 1958, *Publ. Astron. Soc. Japan*, 8, 108.

West, E.A., Hagyard, M.J., 1983, *Solar Phys.*, 88, 51.

Wittmann, A., 1971, *Solar Phys.*, 20, 365.

Ye Shi-hui, Jin Jie-hai, 1983, *Chinese Astron. Astrophys.*, 7, 255.

DYNAMIC MODELS OF FLUX TUBES IN THE
INTERPRETATION OF POLARIZATION MEASUREMENTS

E. RIBES (1), D. REES (2) and FANG CHENG (3)

- (1) DASOP, OBSERVATOIRE DE MEUDON, 92190 MEUDON
- (2) DEPARTMENT OF APPLIED MATHEMATICS, UNIVERSITY OF SYDNEY
- (3) DEPARTMENT OF ASTRONOMY, NANJIN UNIVERSITY.

Recent observations of Stokes parameter profiles indicate the presence of mass motions with large velocity gradients associated with small-scale magnetic elements. Dynamic models of flux tubes have to be used in order to interpret observations of unresolved elements. It is clear that the physical picture of the dynamic models will be quite different from the hydrostatic ones since there is a strong coupling between the magnetic and the velocity field. Self-consistent models should take into account the thermal and dynamic structures simultaneously.

Two-D steady flow solutions in slender magnetic tubes have been worked out (Unno and Ribes, 1979; Webb, 1980). The force driving the flow along the magnetic lines of force is the difference between the gas pressure gradient and the gravity. The heating mechanism is provided by the presence of a downdraft which is carrying an entropy excess from above. The entropy excess Δs between the inside and the outside of the magnetic element is likely to exist since the density inside is lower and the temperature higher ($\Delta S = \ln T_{in}/T_{out} + \gamma \ln \rho_{out}/\rho_{in}$). In the high photosphere (or the chromosphere) where the density is low, the radiative losses are small. As a consequence, the entropy excess causes an enhancement of the brightness of the magnetic element. As the flow goes down, the radiative losses become important (density is higher and higher) until they balance the entropy excess. At that level, the temperature in the tube and out is the same.

For a steady flow, the problem of mass supply has to be considered. However, there is enough solar activity in the form of Evershed flow, surges, spicule material, etc., to overpopulate the chromosphere. This material excess there acts as a pressure inducing a downflow along the magnetic lines of force. From high resolution observations (Muller, private communication), the lifetime of the downdraft seems to be of the same order as the spicule lifetime. If so, the mass supply is not a problem and the assumption of steadiness is reasonable. Moreover, the downdraft brings cooler material down to deeper layers and provides a mechanism for concentrating the flow (Parker, 1978).

1 - Properties of the steady Bernoulli solutions

In a magnetic configuration opening upwards, the Bernoulli solutions have properties similar to those of the solar wind (Parker, 1964): there is a singular point where the flow becomes supersonic. Two critical solutions passing through the singular point have the properties of shock solutions since they show rapid changes of density and velocity. Other families of solutions exist and are determined from the boundary conditions. Among the solutions relevant to the physics of the small-scale fields, one can distinguish 2 classes:

- Those corresponding to moderate field strength (B at the base of the photosphere < 1 kilogauss), and a downdraft decelerating with increasing depth. They exhibit a temperature excess localized in the upper photosphere (Fig. 1, models 1 and 2).

- Those corresponding to a kilogauss field associated with a downdraft accelerating with increasing depth (Fig. 1 models 3 and 4). In model 3, the temperature excess is important throughout the photosphere. This situation is thermally unstable because a strong ΔT localized in structures of less than $1/3''$ will be radiated away in a few seconds. In model 4, the flow is quasi-adiabatic in order to prevent the radiation from escaping. As a result, the flow reaches sonic velocities. A priori, none of these 4 solutions can be disregarded since individual magnetic elements are unresolved. The only way to discriminate among the various Bernoulli solutions is to compare observed line profiles with those derived from the magnetic models.

2 - Observational diagnostics for dynamic models

A model diagnostic should be based on the following set of observational criteria:

- 1) the central intensity of the line
- 2) the continuum excess
- 3) the line weakening
- 4) the shape of the V Stokes profile and the line shift
- 5) the net circular polarization (degree and sign).

We have calculated the theoretical profiles for various photospheric lines with simple Zeeman splitting assuming L.T.E. for a longitudinal magnetic field. The transfer equations are solved at the center of the tube.

Detailed calculations are presented in a forthcoming paper (Ribes et al., 1984). All models predict substantial variation of the central intensity of the line and line weakenings, because they all exhibit temperature excess. However, the asymmetries of the V Stokes profile which are governed by dynamics differ very much from one model to the next (Fig.2). According to recent observations of the Stokes parameters (Bezanger and Semel, 1984; Stenflo et al. 1984), the blue peak amplitude of the V Stokes profile should be larger than the red one. Also, the blue area should be larger than the red area (positive net circular polarization). An important result is that the sign of the net circular polarization is related to the sign of the velocity gradient, and, therefore, is a good way to discriminate between strong field and moderate field solutions. The comparison between the observed V Stokes profile and the theoretical ones (Fig.2) favors moderate field concentration and a downdraft decelerating with increasing depth (model 2).

One point remains unclear, that is the small apparent redshift of the v zero crossing point found by all observers. Strong velocity gradients associated with an updraft or a downdraft should shift the line significantly. However, a downflow is more likely to exist since it provides a mechanism for concentrating the magnetic field.

When the field strength inside the tube is large enough, the flow may exhibit a stable oscillatory behaviour (Hasan, 1984). With a lack of spatial, spectral or temporal resolution, no appreciable downflow in flux tubes will be observed. In Hasan's model, the instability (downflow) persists for 10 minutes which, again, makes the steady approximation correct.

Association of spicules and photospheric magnetic fields might be another explanation of the small apparent redshift (Parker, private communication). Some eruption of gas moving upwards occurs in the tube and supplies a reaction downflow. If so, both updraft and downdraft could be present in the magnetic observations reducing the net line shift.

CONCLUSION

Dynamic models lead to solutions quite different from hydrostatic ones. In view of the detection of mass motions with large velocity gradients, polarization measurements have to be interpreted in terms of dynamic models.

The main properties of the intensity line profiles as well as the asymmetries of the V Stokes profiles can be explained best in terms of magnetic elements with moderate field strength.

REFERENCES

- BEZANGER C., SEMEL M.: 1984, to be submitted
HASAN S.S.: 1984, *Astrophys. J.*, in press.
PARKER E.N.: 1964, *Astrophys. J.* 140, 1170.
PARKER E.N.: 1978, *Astrophys. J.* 221, 368.
RIBES E., REES D. and FANG CHENG: 1984, *Astron. and Astrophys.*, submitted.
STENFLO O., HARVEY J.W., BRAULT J.W. and SOLANSKI S.: 1984, *Astron. and Astrophys.*, in press.
UNNO W., RIBES E.: 1979, *Astron. and Astrophys.* 73, 114.
WEBB A.R.: 1980, Ph.D. Thesis, St. Andrews University.

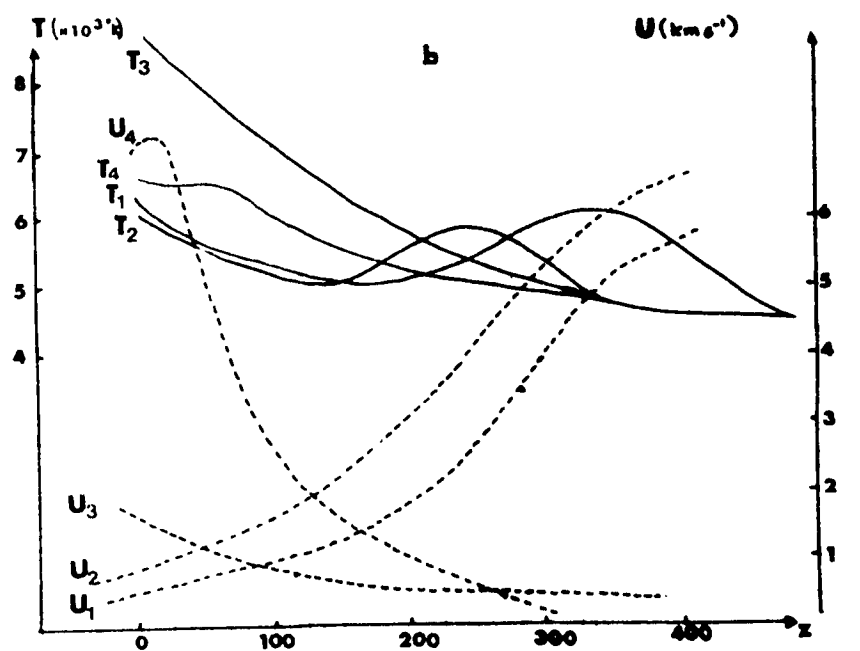
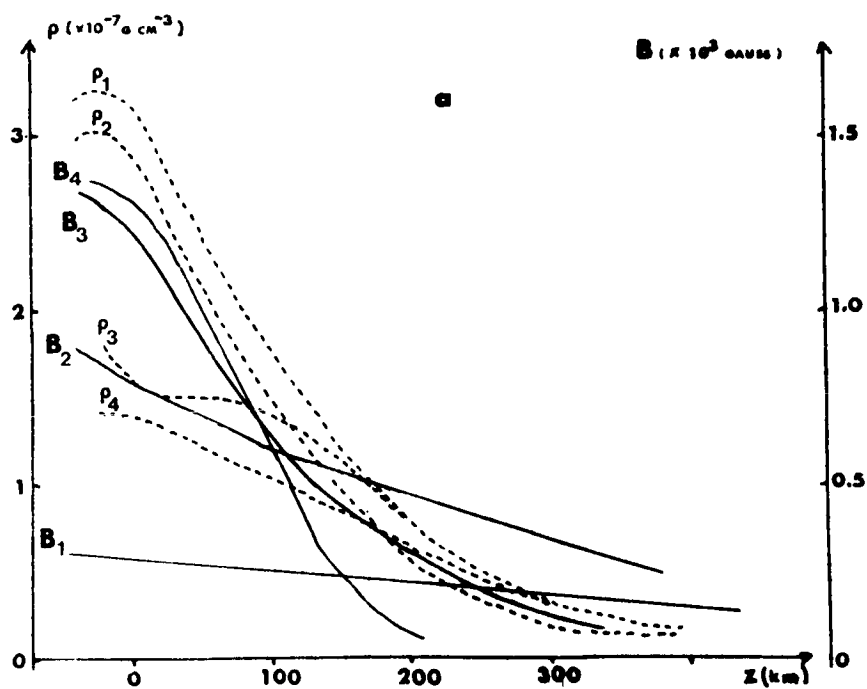


Figure 1 - Magnetic field, velocity field, density and temperature are given as functions of height for models 1 to 4.

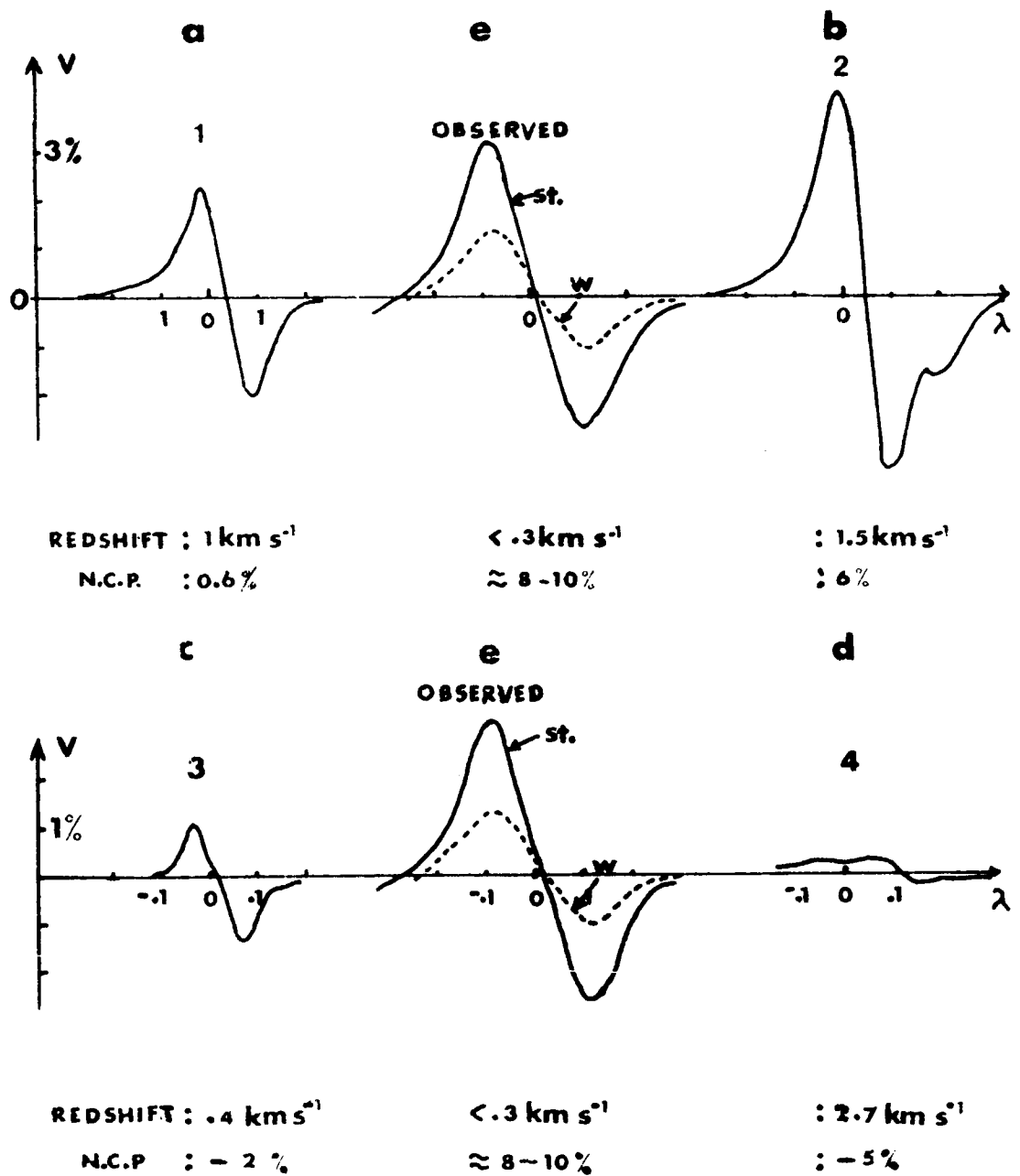


Figure 2 - V Stokes profiles of the 5250.22 λ line have been calculated for models 1 to 4 (2 a.b.c.d.). The V profile observed by Stenflo et al. (1983) is given in 2e where full and dashed lines correspond to strong and weak plages respectively.

Note that the sign of net circular polarization is positive only for models 1 and 2.

LEAST SQUARES INVERSION OF STOKES PROFILES
IN THE PRESENCE OF VELOCITY GRADIENTS

A. SKUMANICH

High Altitude Observatory, NCAR¹, Boulder, CO. 80307, U.S.A.

D. E. REES

Dept. of Applied Mathematics, Uni. of Sydney, N.S.W. 2006, Australia.

and

B. W. LITES

National Solar Observatory, NOAO², Sunspot, N.M. 88349, U.S.A.

Abstract. The Auer, Heasley and House Stokes inversion procedure in use at HAO is based on an analytic solution of the equation of transfer for polarized light where the representation of the thermodynamic and magnetic structure of the atmosphere is assumed to have a high degree of 'invariance', namely, a Milne-Eddington (ME) structure with a constant magnetic field. In the presence of invariance breaking gradients the resultant Stokes profiles are represented only approximately, if at all, by analytic forms. The accuracy of the inversion parameters and their significance as 'measures' of actual structure are explored for the ME and the Landman-Finn sunspot models under the effects of velocity gradients. The resulting field parameters are good to a few percent and prove to be insensitive to the errors committed by the use of a ME-representation, but the resulting ME parameters yield a less precise measure of thermal structure. In particular η_0 , the line to continuum opacity at line centre, can be significantly underestimated (as much as a factor of 9) while the Doppler width is overestimated by as much as 20%.

¹ The National Center for Atmospheric Research is operated by the University Corporation for Atmospheric Research under sponsorship of the U.S. National Science Foundation.

² The National Optical Astronomy Observatories are operated by the Association of Universities for Research in Astronomy, Inc. under contract AST 78-17292 with the U.S. National Science Foundation.

1. Introduction

Considerable progress has been made in the development of the theory of spectral line formation in a magnetic field (see review by Landi degl'Innocenti (1984)). Nevertheless the elementary analytic solutions of the transfer equations for a Milne-Eddington (ME) atmosphere and constant magnetic vector (Unno, 1956; Rachkovsky; 1962, 1967) still retain a central role in the interpretation of Stokes parameter profiles. An efficient least squares method for fitting data by adjusting the parameters of this model was devised by Auer, Heasley and House (1977). This was used by Gurman and House (1979) to map the vector magnetic field in a sunspot observed in the Zeeman triplet Fe I 6302.5 Å with the HAO Stokes Polarimeter. Several difficulties were encountered in the application of Auer et al's method (HAO Report, 1982). Landolfi and Landi degl'Innocenti (1982) realised it was essential to include magneto-optical effects in the inversion procedure especially to ensure a reliable estimate of magnetic field azimuth. This and other improvements have been incorporated by Lites and Skumanich (1984a) in an up-dated version of the Auer et al code. With theoretical profiles generated with the Unno-Rachkovsky solutions the new version recovers the input parameters of the ME atmosphere and the magnetic vector to machine accuracy. The degradation of this type of fit by limited spectral resolution is studied by Lites and Skumanich (1984b).

In this paper we study the performance of the least squares fit when applied to theoretical profiles calculated with a realistic model of a sunspot. The model has gradients in all physical parameters including magnetic and velocity fields. This numerical experiment allows one to assess the physical significance of the ME and mean field parameters provided by the inversion. Observed Stokes profiles invariably have asymmetric distortions due to velocity gradients in the solar atmosphere. The inversion method symmetrizes the profiles prior to fitting and it is important to know how this affects the inference on the magnetic vector.

2. Sunspot Model

The calculations are based on the sunspot model of Landman and Finn (1979), henceforth denoted by LF. This model is cylindrically symmetric with a magnetic field structure derived from the 'similarity principle' of Schlüter and Temesvary (1958). More elaborate models are available, e.g. the 'return flux' model (Flaa et al, 1982), but we have chosen the LF model for this initial study as part of a broader investigation on net circular polarization, a topic which motivated the work of LF.

In order to produce an emergent continuum flux consistent with observations across sunspots LF introduced an empirically determined turbulent pressure term into the equation of state. In their line transfer calculations another Evershed-type flow was imposed along the magnetic field lines. These two velocity fields were not physically consistent and LF indicated the need for improvements in the modelling. Indeed a detailed MHD model of a sunspot, with an adequate treatment of the difficult problem of energy transport, remains a goal with high priority. In the spirit of LF's approach we use the LF thermodynamic structure which includes the

effect of the turbulent velocity distribution. However Skumanich (1984) found that their ordered velocity field was not able to explain the net circular polarization data of Henson and Kemp (1984), a more satisfactory velocity model being that of Auer and Heasley (1978). We use the latter which assumes an accelerating upflow along the field lines with the flow velocity at all radii from the centre of the spot being expressed as

$$v = -2/(1 + 5\tau_c)$$

where v is units of the Doppler width and τ_c is the continuum optical depth.

3. Line Formation and Inversion

Theoretical profiles of the Stokes parameters (I,Q,U,V) were calculated for normal emergence from the spot model at selected radii. The transfer equations including magneto-optics were solved using the finite difference method developed by LF (see Skumanich (1984) concerning corrections to this method).

The least squares inversion method has 8 fitting parameters:

(a) The thermodynamic parameters -

η_0 = the line centre to continuum opacity (for zero damping);

B_1 (or β) = the slope of the source function which is represented as a linear function of optical depth τ_c (i.e. $B_V(T_e) = B_0 + B_1\tau_c$ with $B_1 = B_0\beta$);

$\Delta\lambda_D$ = the Doppler width;

a = the Voigt damping width in Doppler width units.

(b) The field parameters -

B = the magnitude of the magnetic field;

ψ = the field inclination to the line of sight;

ϕ = the field azimuth;

$\Delta\lambda_V$ = the position of line centre.

The position $\Delta\lambda_V$ is determined from the wavelength shift of the centroid of the intensity line depth profile. In agreement with Auer et al we found this to be the most reliable point about which to symmetrize the profiles. This corresponds to the point which halves the equivalent width of the split line. Several neighbouring wavelength points were tested on either side of $\Delta\lambda_V$ and the results

indicate that $\Delta\lambda_v$ is in fact the optimal choice in the sense that it gives the smallest value for the χ^2 of the fit. Typically we fit about 42 points at 10 mA steps in the symmetrized profiles.

The magnetic field viewed along a vertical path through the spot model has a constant azimuth. The Stokes profiles are computed in a reference frame X-Y where this azimuth is zero; i.e. the X-axis is along the spot radius (c.f. Figure 1).

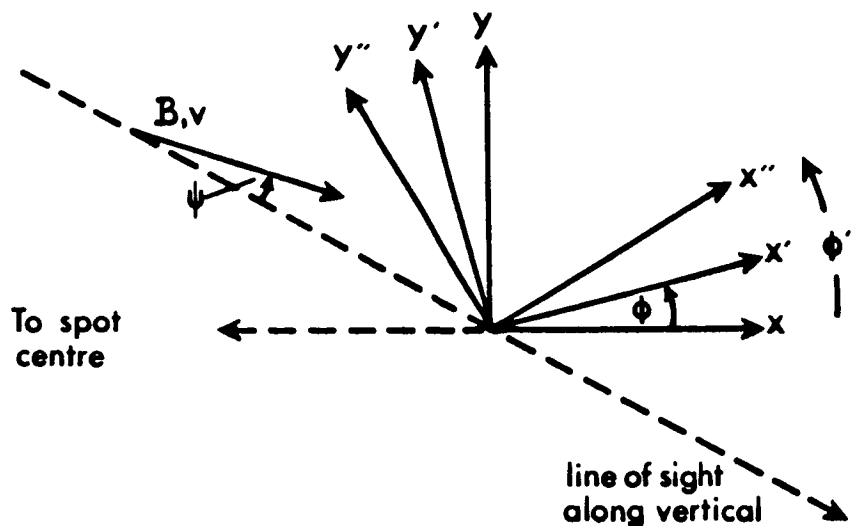


Fig. 1. Profiles are calculated in the canonical frame X-Y. The magnetic field B is at an angle ψ to the line of sight and has zero azimuth relative to X-Y; the velocity field v is parallel to B. X''-Y'' is the frame in which ϕ'' is interpreted as the field azimuth, magneto-optics being neglected. X'-Y' is the pseudo-canonical frame, ϕ being the azimuth of the fit including magneto-optics.

We call this the canonical frame. In real data such a canonical frame will not exist because of variations in field azimuth along the line of sight. At best one could search for an approximation to the canonical frame which in some sense would refer to the average field azimuth.

In the canonical frame the computed Stokes parameter U is non-zero owing to magneto-optical effects; it would be zero in their absence. It is instructive to rotate the "data" (I,Q,U,V) through an angle ϕ'' to a new frame X''-Y''. The Stokes vector in this frame is (I'',Q'',U'',V'') where, in particular,

$$U'' = U \cos 2\phi'' - Q \sin 2\phi''$$

Auer et al chose ϕ'' so as to minimise $\Sigma(U'')^2$, the sum being over the wavelength points sampled in the profile. Then they ignored magneto-optics, interpreted ϕ'' as the field azimuth and restricted the fit to the Stokes parameters I'' , Q'' and V'' . Here we fit U'' as well, allowing the azimuth to be a free parameter. Thus we obtain a better estimate of the field azimuth ϕ relative to the canonical frame. To be consistent with the way one would handle real observations we present graphs of the "data" and the fits in the frame $X'-Y'$ (c.f. Figure 1) which is rotated through ϕ relative to the $X-Y$ frame. Since the model field azimuth is zero, ϕ is the azimuth error in the inversion; the $\pm 180^\circ$ ambiguity is, of course, ever present. We call $X'-Y'$ the pseudo-canonical frame.

4. Field Gradient Effects

As noted above, the inversion recovers to machine accuracy the ME and constant magnetic field parameters from profiles generated with the Unno-Rachkovsky solutions. Prior to treating the fully variable sunspot model we test the effects of a velocity gradient alone, and then a combination of velocity and magnetic field gradients. The profiles are computed with ME parameters $\eta_0 = 20$ and $\beta = 9$ for a Zeeman triplet with splitting equivalent to a line centred at wavelength 5303 \AA with Landé factor $g_J = 3$. The Doppler width $\Delta\lambda_D = 35 \text{ m\AA}$ and the damping width $a = 10^{-4}$ (i.e. essentially a Gaussian absorption profile).

(a) Velocity Gradient

The constant magnetic field input values are taken here to be averages over the continuum optical depth range $\tau_c = .05$ to $.5$ in the LF model. In Table I we present the parameter fits for a sequence of radial positions r/r_p where $r_p = 2.2 \times 10^9 \text{ cm}$ is the penumbral radius for the LF model (see Skumanich, 1984).

The field strength B and inclination ψ are found to be determined to within 2%. Although significant errors in azimuth do occur for small ψ values (e.g. $\phi = -4.3^\circ$ at $\psi = 4^\circ$), in most cases the symmetry breaking by the velocity gradient has a negligible effect on the inferred magnetic field.

The error in the source function slope β is less than 3%, but the other ME parameter η_0 is systematically underestimated by 15-20%, a result which heralds difficulties for a meaningful interpretation of thermodynamic parameters derived from real data. The Doppler width is overestimated by 8-21%, the error decreasing as the field inclination ψ increases. As expected, the velocity gradient induced asymmetries are interpreted as additional random broadening in the symmetrized profiles. The damping width is in error by as much as a factor of 27 but still takes values typical of approximately Gaussian absorption profiles.

TABLE I

Parameter fits for ME model

 $(\eta_0 = 20, \beta = 9, \Delta\lambda_D^0 = 35 \text{ mÅ}, a = 10^{-4})$ with a velocity gradient.

r/r_p^\dagger	.05	.11	.23	.45	.68	.91
$B(\text{G})$	3433	3370	3113	2223	1236	498
$B^f(\text{G})^*$	3432	3367	3100	2183	1214	498
$\psi (\text{deg})$	4.0	9.0	17.8	33.0	43.5	51.7
$\psi^f (\text{deg})$	4.1	9.2	17.8	32.4	43.2	52.1
$\phi^f (\text{deg})$	-5.0	-4.3	-2.6	-0.6	0.3	0.5
β^f	8.73	8.74	8.79	9.04	9.18	9.33
η_0^f	16.70	17.05	17.51	17.04	17.55	15.67
$\Delta\lambda_D^f (\text{mÅ})$	42.5	42.5	41.8	40.8	38.8	37.9
$a^f (\times 10^4)$	1.0	1.0	1.1	0.5	2.8	27.0

 $\dagger r_p = 2.2 \times 10^9 \text{ cm}$ is the penumbral radius

* Here and in Tables II and III fits are indicated by the superscript f

(b) Velocity and Magnetic field Gradients

We now include gradients in B and ψ provided by the LF model. The inversion returns only a pair of values (B, ψ) which must be related to a particular depth or range of depths in the atmosphere. This raises the complex issue of the depth of formation of magnetically sensitive lines. To highlight the problem, consider a Zeeman triplet formed in a purely longitudinal magnetic field. Suppose we observe at line centre. As the field strength increases, the total opacity

for the Stokes parameter I decreases and we see deeper into the atmosphere. In the limit where the sigma components are completely separated, the total opacity at line centre is equal to that in the continuum! Alternatively, if τ_0 denotes the total optical depth at line centre, then $\tau_0 \rightarrow \tau_c$ as $B \rightarrow \infty$. Several authors have considered methods of assigning depths of formation to the Stokes parameters (Rachkovsky, 1969; Staude, 1972; Landi degl'Innocenti and Landi degl'Innocenti, 1977) but their results are inconclusive. A promising new approach discussed by van Ballegooijen (1984).

Despite the inherent ambiguities we adopt here the simple expedient of comparing the fits (B, ψ) with values at $\tau_0 = 2/3$.

The fits are shown in Table II. The errors in all parameters are similar to those cited in the previous section, except at $r/r_p = .05$ where $\eta_0 = 8.44$, less than 50% the input value of 20. The large azimuth error $\phi = -13.3^\circ$ is another indication of the poor fit one can expect at small ψ values.

TABLE II

Parameter fits for ME model ($\eta_0 = 20$, $\beta = 9$, $\Delta\lambda_D^0 = 35 \text{ m}\text{\AA}$, $a = 10^{-4}$)
with velocity and magnetic field gradients.

r/r_p	.05	.11	.23	.45	.68	.91
$B(G)$	3442	3386	3109	2230	1271	560
$B^f(G)$	3436	3370	3078	2188	1247	557
$\psi \text{ (deg)}$	4.0	9.0	17.8	33.0	44.0	52.0
$\psi^f \text{ (deg)}$	5.3	9.2	17.8	32.4	43.8	52.5
$\phi^f \text{ (deg)}$	-13.3	-4.2	-2.6	-0.8	+0.3	+0.4
β^f	9.10	9.01	8.98	9.04	9.19	9.50
η_0^f	8.44	17.40	17.68	16.85	17.50	15.80
$\Delta\lambda_D^f \text{ (m}\text{\AA})$	50.1	42.5	41.9	41.0	38.8	37.9
$a^f (\times 10^4)$	1.0	0.1	1.0	1.0	1.6	32.

5. Fully Variable Spot Model Calculations

Section 4 focused on field gradients without the confusion of gradients in atmospheric conditions affecting η_0 , the source function and the Doppler width. We add this new level of complexity by computing Stokes profiles for a Fe I line of the type considered by LF. The data for the LTE calculations are: iron abundance relative to hydrogen, $10^{-4.4}$; ionization potential, 5.25 eV; excitation potential, 3.28 eV; central wavelength, 5303 Å; oscillator strength, $gf = 1$; Landé factor, $g_J = 3$. Lines with different excitations, strengths and splittings were also tested but details are not given since they gave similar results.

Figure 2 shows profiles in the pseudo-canonical frame ($X'-Y'$) at the radial position $r/r_p = .23$ where the fits to (B, ψ, ϕ) are $(3058 \text{ G}, 17.6^\circ, 1.1^\circ)$. The model profiles are the full lines and they exhibit asymmetries common to all our calculations. In particular the V profile has a blue shift and a blue enhanced asymmetry, though the latter feature is not evident on the scale of the figure. The implications of such properties of V for the study of net circular polarization are discussed by Skumanich (1984) and Ribes et al (1984). The dashed lines are the fits to the symmetrized profiles. Note that the U Stokes parameter is positive at all wavelengths. This is an indication of the fact that the field azimuth is indeed estimated accurately, the pseudo-canonical frame almost coinciding with the canonical frame.

The wavelength shift $\Delta\lambda_V$ of the symmetrized profiles provides an estimate of the average velocity field. This shift is derived by locating the centroid of the intensity line depth profile. An alternative velocity measure is the shift of the zero-crossing of the V profile (e.g. Giovanelli, 1970). In all cases we found the two shifts to coincide, and this is clear in Figure 2. However there is a significant displacement between the apparent centre of the model Q profile and the centre of the Q fit (the latter of course has a shift $\Delta\lambda_V$). We shall not detail specific values here, but simply note that velocities inferred from shifts of the bisector of the Q peaks for example are systematically greater than those obtained from $\Delta\lambda_V$. The velocity increases with height in the sunspot model and such differences may indicate that the depth of formation for Q is separated from that for I and V. There could be a possible way of extracting velocity gradients from such comparisons. One important point has emerged from preliminary investigation of this phenomenon: 10 mÅ sampling of the profiles is not always adequate to resolve accurately the Q shift from the (I,V) shift, and one must sample at smaller intervals, say 5 mÅ. To decide whether or not it is warranted to pursue this potentially new diagnostic, with its demand for high spectral resolution, will require more detailed modelling studies.

Profiles referred to the frame $X''-Y''$ are shown in Figure 3 for $r/r_p = .34$. The angle of rotation is $\phi'' = 11.6^\circ$, which would be the assumed field azimuth if magneto-optics were neglected. Note that at wavelengths near line centre the Stokes parameter U'' is negative. Such negative values in real data might be regarded as pathological, but they are simply the result of "cross-talk" between

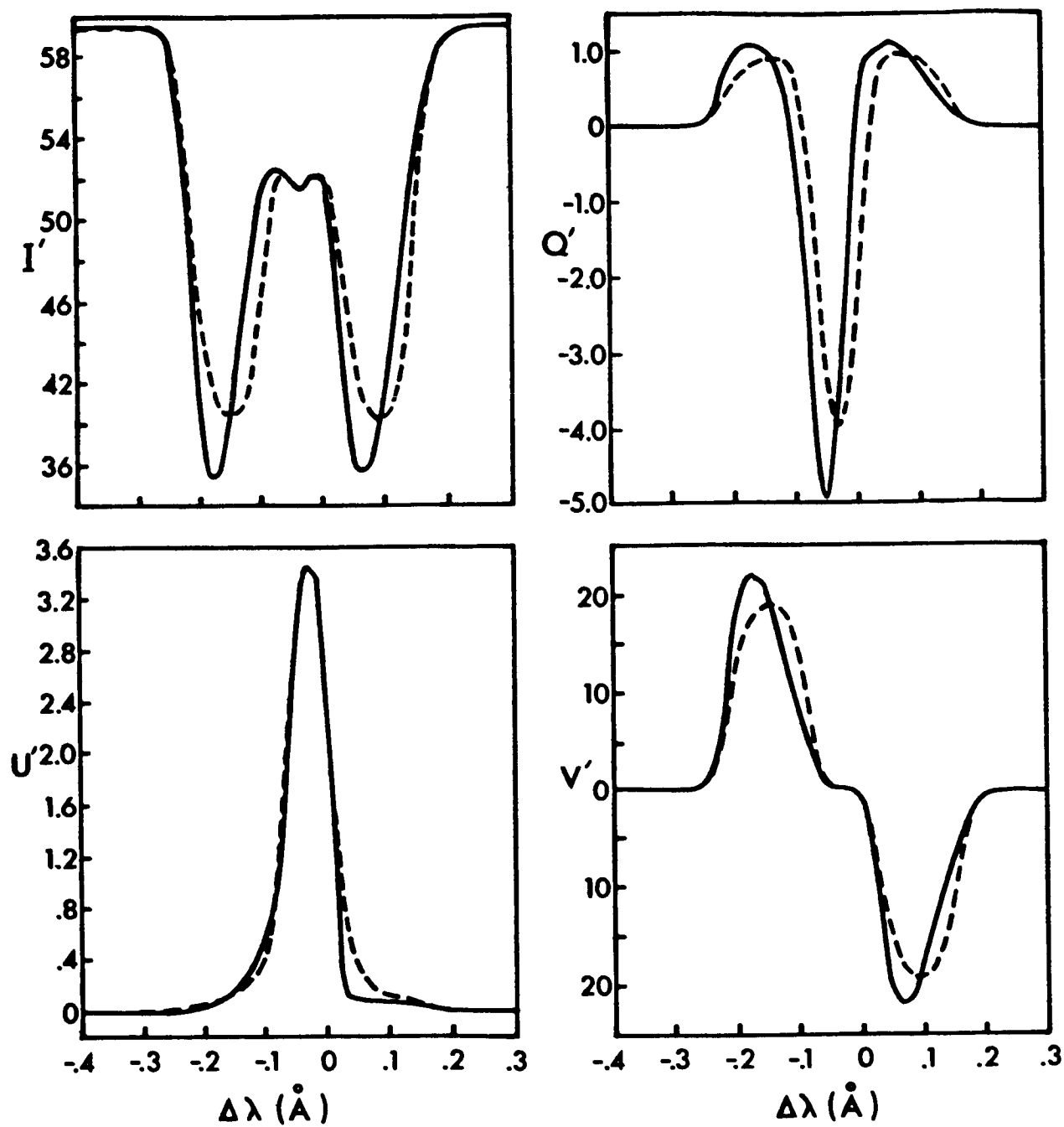


Fig. 2. Stokes profiles (— model calculations; ---- symmetrized fits) in the pseudo-canonical frame $X'-Y'$ for the LF model at $r/r_p = .23$. (I', Q', U', V') are in units of $10^{-7} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \text{ Hz}^{-1}$.

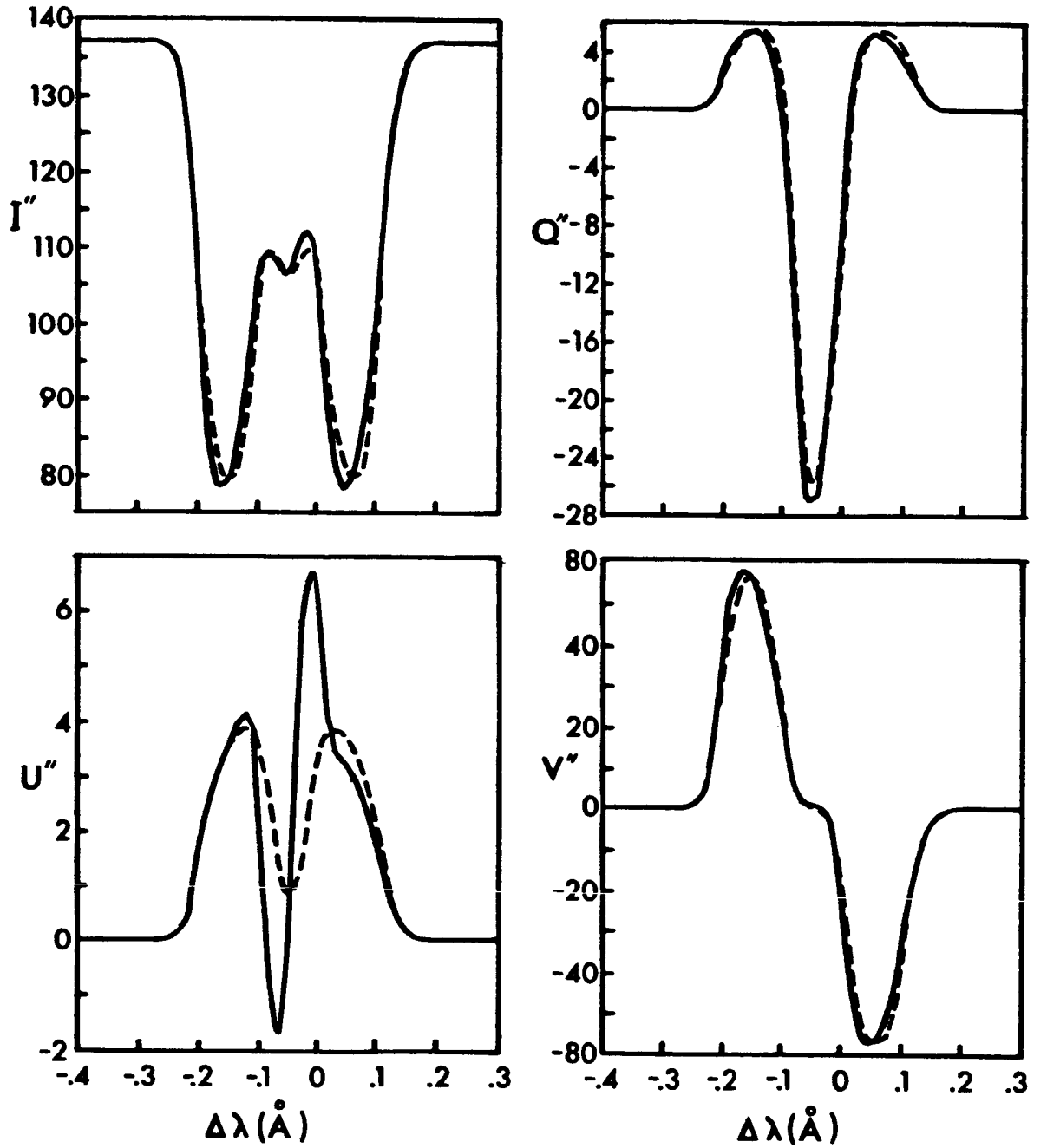


Fig. 3. Stokes profiles (— model calculations; --- symmetrized fits) in the frame X'' - Y'' for the LF model at $r/r_p = .34$. Note the negative values of U'' due to the large rotation angle ϕ''_p relative to the canonical frame. (I'' , Q'' , U'' , V'') are in units of 10^{-7} erg cm $^{-2}$ s $^{-1}$ sr $^{-1}$ Hz $^{-1}$.

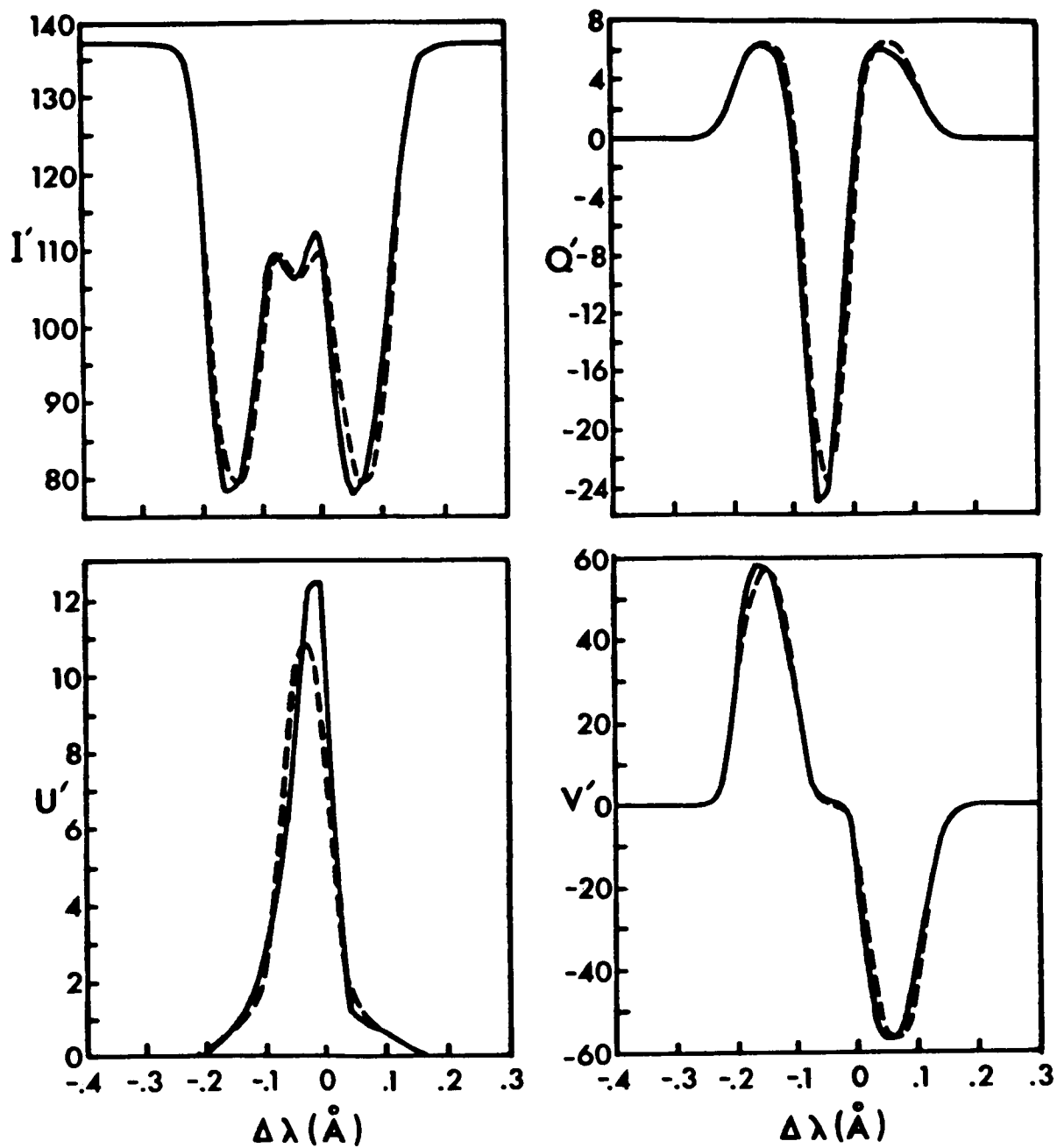


Fig. 4. The same Stokes profiles as in Fig. 3. relative to the pseudo-canonical frame $X'-Y'$.

the Q and U linear polarizations. When rotated to the pseudo-canonical frame (Figure 4) the profile U' is positive at all wavelengths, as expected from the small azimuth error $\phi = .2^\circ$ obtained at this radial position. The parameters Q'' and Q' in the respective frames differ only slightly, while $I' = I''$ and $V' = V''$ since these two Stokes parameters are invariant under rotation of reference frame.

Details of the fits are given in Table III. Because of problems with the thermodynamic structure near the axis of the LF model (see discussion by Skumanich (1984)), we have restricted the analysis to $r/r_p \geq .23$. Except for the source function slope B_1 , the parameters are compared with model values at $\tau_0 = 2/3$. As in the ME model results, the magnetic field vector is very well determined (differences of less than a few percent for B and ψ ; and $|\phi| \lesssim 2^\circ$). Also the slope B_1 is closely correlated with the average slope from $\tau_c = .05$ to $.5$ (differences ranging over 6-38%). The Doppler width is overestimated by 5-22%, while the damping width is severely overestimated at large values of ψ (e.g. by a factor of 135 at $\psi \approx 52^\circ$).

It may be too restrictive to compare the η_0 fit with a value at a single depth in the atmosphere, but the huge differences (underestimates by factors as high as 9) are not atypical over the range of depths where the line is formed. Clearly η_0 is a convenient intermediate parameter leading, in particular, to a good estimate of the magnetic vector. However its value as a diagnostic of thermodynamic structure must be seriously questioned.

Simultaneous observation of magnetically sensitive and insensitive lines is often suggested as way to disentangle magnetic structure from thermodynamic structure. In this context the ambiguities associated with the fit to η_0 can be rather well illustrated thus. Consider two situations, one where the Fe I line has a Landé factor $g_J = 0$, the other where $g_J = 3$. Inversion of profiles calculated at $r/r_p = 0.45$, for example, give the following η_0 -fits: for $g_J = 0$, $\eta_0 = 6.9$; for $g_J = 3$, $\eta_0 = 4.5$. The other parameters (B_1 , $\Delta\lambda_D$, a) were not significantly different in the two cases. The important point is that the η_0 of the magnetically insensitive line exceeds that of the sensitive line by over 50%. This implies that any attempt to calibrate the thermodynamic parameters of a sensitive line using those from an insensitive line must be treated with caution.

TABLE III

Parameter fits for fully variable sunspot model
and archetypal Fe I line at 5303\AA°

r/r_p	.23	.34	.45	.57	.68	.80	.91
$B(G)$	3120	2701	2230	1738	1269	872	563
$B^f(G)$	3058	2678	2228	1737	1271	870	561
ψ (deg)	17.7	25.9	33.0	39.0	44.0	48.5	52.0
ψ^f (deg)	17.6	26.9	34.7	38.9	44.3	48.1	52.2
ϕ^f (deg)	1.1	0.2	-1.8	-2.1	-1.1	-0.9	-0.5
B_1^*	0.76	2.14	2.09	1.38	2.62	4.32	4.03
B_1^f	0.47	1.63	1.96	1.52	1.99	3.13	3.64
η_0	13.0	11.7	20.6	23.3	27.9	18.7	8.8
η_0^f	4.9	2.7	3.8	6.1	2.9	2.1	1.4
$\Delta\lambda_D^{\circ}$ (mÅ)	36.2	36.8	35.9	35.6	35.5	36.3	37.1
$\Delta\lambda_D^f$ (mÅ)	44.8	41.9	41.1	40.0	39.8	38.6	39.8
a^f ($\times 10^4$)	48	1	1	.4	18	103	135

* B_1 and B_1^f are in units of $10^{-5} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \text{ Hz}^{-1}$

6. Conclusion

The HAO computer code for least squares fitting of Stokes profiles is a powerful and reliable diagnostic tool, even though it is limited by the assumptions of a static Milne-Eddington model atmosphere and constant magnetic vector. In this study profiles distorted by velocity gradients were symmetrized prior to fitting, though it is possible to fit the entire asymmetric profile (Lites and Skumanich, 1984b). Our model calculations show that this symmetrization has negligible effect on the quality of the fit to the magnetic field parameters. In most cases the field strength B and inclination ψ to the line of sight are determined to within a few percent. The field azimuth ϕ is accurate to $\sim 1\text{--}2^\circ$ except when the field is nearly longitudinal; then large errors $\gtrsim 10^\circ$ can occur.

The symmetrization introduces an artificial random broadening into the profiles leading to an overestimate of the Doppler width of up to $\sim 20\%$. Since we have restricted the calculations to essentially Gaussian absorption profiles it is not possible to comment definitively on the role of the damping width. The source function slope agrees well with average values, always being within a factor of 2.

The most significant result for the thermodynamic structure is the extremely poor fit to η_0 , the line to continuum opacity at line centre. It can be underestimated by as much as a factor of 9 when comparisons are made at particular depths in the atmosphere.

These conclusions must be tempered by the realisation that it is very difficult to decide on an optimal way to assign a depth of formation for a magnetically split line. Perhaps van Ballegooijen's (1984) approach may be useful here.

Subject to this caveat the general impression is that one can confidently derive the magnetic vector from Stokes profiles (at least of lines formed in LTE), but that the thermodynamic structure in a magnetic region is far more elusive. Note that we have ignored several important effects, e.g. unresolved fine structure, scattered light and instrumental noise. These may degrade the fits to both magnetic and thermodynamic structure, but are not expected to affect the relative performance of the inversion, i.e. the derived magnetic structure should still be the more reliable.

No doubt it will be necessary to do detailed LTE and non-LTE modelling along with profile inversions in order to try to resolve these issues as they arise in real data. To this end we are developing a "user-friendly" computer code for general non-LTE transfer calculations in a magnetic field (Durrant,

et al, 1984). At the same time it is important to extend the inversion method to source functions with a non-linear dependence on optical depth so that strong chromospheric lines may be analysed.

Finally, we have hinted at the possibility of using relative displacements of the V and Q profiles to infer velocity gradients. Again this demands a thorough investigation of the theory of depth of formation of Stokes profiles.

Acknowledgement

This work was supported in part by an ARGS grant for the project "Stokes Polarimetry and Sunspot Chromospheres" and in part by funds from the University of Hawaii. We thank C. Durrant for her invaluable computing assistance.

References

- Auer, L. H. and Heasley, J. N.: 1978, *Astron. Astrophys.* **64**, 67.
- Auer, L. H., Heasley, J. N. and House, L. L.: 1977, *Solar Phys.* **55**, 47.
- Durrant, C. J., Rees, D. E. and Skumanich, A.: 1984, in preparation.
- Flaa, T., Osherovich, V. and Skumanich, A.: 1982, *Astrophys. J.* **261**, 700.
- Giovanelli, R. G.: 1970, *Proc. Astron. Soc. Australia* **1**, 363.
- Gurman, J. and House, L. L.: 1979, *Solar Phys.* **71**, 5.
- HAO Report: 1982, *Bull. Am. Astron. Soc.* **14**, 195.
- Henson, G. D. and Kemp, J. C.: 1984, submitted to *Solar Phys.*
- Landi degl'Innocenti, E. and Landi degl'Innocenti, M.: 1977, *Astron. Astrophys.* **56**, 111.
- Landi degl'Innocenti, E.: 1984, in these proceedings.
- Landman, D. A. and Finn, G. D.: 1979, *Solar Phys.* **63**, 221.
- Landolfi, M. and Landi degl'Innocenti, E.: 1982, *Solar Phys.* **78**, 355.

- Lites, B. W. and Skumanich, A.: 1984a, SPO Internal Memorandum T2-84.
- Lites, B. W. and Skumanich, A.: 1984b, in these proceedings.
- Rachkovsky, D. N.: 1962, *Izv. Krymsk. Astrofiz. Observ.* 27, 148.
- Rachkovsky, D. N.: 1967, *Izv. Krymsk. Astrofiz. Observ.* 37, 56.
- Rachkovsky, D. N.: 1969, *Izv. Krymsk. Astrofiz. Observ.* 40, 127.
- Schlüter, A. and Temesvary, S.: 1958, in E. Lehnert (ed.), 'Electromagnetic Phenomena in Cosmical Physics', *IAU Symp.* 6, 263.
- Skumanich, A.: 1984, submitted to *Astrophys. J.*
- Staude, J.: 1972, *Solar Phys.* 24, 255.
- Unno, W.: 1956, *Publ. Astron. Soc. Japan* 8, 108.
- Van Ballegooijen, A.: 1984, in these proceedings.

CONTRIBUTION FUNCTIONS FOR ZEEMAN-SPLIT LINES, AND LINE FORMATION IN PHOTOSPHERIC FACULAE

A.A. van Ballegooijen
Lockheed Palo Alto Research Laboratory

ABSTRACT

The transfer of polarized light in an inhomogeneous stellar atmosphere, and the formation of magnetically sensitive spectral lines, are discussed. A new method for the solution of the transfer equations is proposed. The method gives a natural definition of the contribution functions for Stokes' parameters, i.e. functions describing the contributions from different parts along the line-of-sight (LOS). The formalism includes all magneto-optical effects, and allows for an arbitrary variation of magnetic field, velocity field, temperature, density, etc. along the LOS.

To illustrate the method I discuss the formation of FeI λ 5250.2 in photospheric faculae. A potential-field model of a facular element is presented, and spectral profiles and contribution functions are computed for the Stokes parameters I, Q, and V.

1. INTRODUCTION

The use of spectral lines to determine the thermodynamic and magnetic structure of the solar atmosphere often requires knowledge of the position along the LOS from where the light, observed at a certain wavelength, originates. For non-magnetic lines, the mean height-of-formation can be computed from a model of the atmosphere, since the emergent intensity at each wavelength can be expressed as an integral over height h of the contribution function:

$$C_I(h) = \frac{1}{\mu} \{ \kappa_c(h) B(h) + \kappa_\ell(h) S(h) \} \exp[-\tau(h)]. \quad (1)$$

Here $B(h)$ is the Planck function, $S(h)$ is the line source function, $\kappa_\ell(h)$ and $\kappa_c(h)$ are the line- and continuum opacity (per unit length), $\tau(h)$ is the sum of line- and continuum optical depth as measured along the LOS, and $\mu = \cos \theta$ defines the angle with the vertical direction.

It is not obvious how to extend this definition to magnetically sensitive lines. In this case the variations of the four Stokes parameters I, Q, U, and V along the LOS are described by four coupled differential equations (e.g. Landi Degl'Innocenti and Landi Degl'Innocenti, 1972). In general these couplings seem to prevent us from writing the emergent intensities as simple integrals over height (for an iterative solution see Staude, 1969). Analytical solutions to the transfer equations can be obtained only for simple models of the line-formation process, in which the magnetic field is either weak (Landi Degl'Innocenti and Landi Degl'Innocenti, 1973), or nearly constant with height (Unno, 1956; Kjeldseth-Moe, 1968; Landi Degl'Innocenti and Landi Degl'Innocenti, 1977). These models do not give a good description of the strong, inhomogeneous magnetic field present on the sun. For more realistic models of magnetic- and atmospheric structure the radiative transfer equations must be solved by numerical integration (Beckers, 1969; Stenflo, 1971; Wittmann, 1974, 1977; Landi Degl'Innocenti, 1976). A disadvantage of the numerical method is, however, that all information about contribution functions and height-of-formation is lost.

In Sect. 2 I show how the four coupled transfer-equations may be formally integrated, leading to a general definition of the contribution function for each Stokes parameter. The solution is not restricted to certain magnetic- or atmospheric models, and includes magneto-optical effects.

In Sect. 3 I discuss the application of my method to a model of photospheric faculae.

2. FORMAL SOLUTION OF TRANSFER EQUATIONS

The transport of polarized light in a magnetically sensitive spectral line is described by the matrix equation:

$$\frac{d}{d\tau_c} \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \begin{bmatrix} (1+\eta_I) & \eta_Q & \eta_U & \eta_V \\ \eta_Q & (1+\eta_I) & \rho_V & -\rho_U \\ \eta_U & -\rho_V & (1+\eta_I) & \rho_Q \\ \eta_V & \rho_U & -\rho_Q & (1+\eta_I) \end{bmatrix} \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} - \begin{bmatrix} B+\eta_I S \\ \eta_Q S \\ \eta_U S \\ \eta_V S \end{bmatrix}, \quad (2)$$

where $I(\tau_c)$, $Q(\tau_c)$, $U(\tau_c)$, and $V(\tau_c)$ are the Stokes parameters at a certain wavelength in the line, and τ_c is the continuum optical depth as measured along the line of sight. The quantities η_I , η_Q , η_U , and η_V are ratio's of the line opacity in each Stokes parameter and the continuum opacity, while ρ_Q , ρ_U , and ρ_V describe

magneto-optical effects. The η 's and ρ 's are functions of continuum optical depth τ_c , and of wavelength in the line; they can be computed if the variations of magnetic field, temperature, density, etc. along the line of sight are known (e.g. Landi Degl'Innocenti, 1976). The emergent intensities $I(0)$, $Q(0)$, $U(0)$, and $V(0)$ can be computed by integrating Eq.(2) numerically. The integration starts at a large optical depth, and proceeds towards smaller depth.

The symmetries and anti-symmetries in the real-valued 4×4 -matrix of Eq.(1) represent a certain redundancy in the formalism. This redundancy is removed by using Jones-calculus, in which the polarization state of the light, and the absorption- and emission properties of the medium, are described by complex 2×2 matrices. I define the matrices $D(\tau_c)$, $A(\tau_c)$ and $F(\tau_c)$ by:

$$D = \frac{1}{2} \begin{bmatrix} I+Q & U+iV \\ U-iV & I-Q \end{bmatrix}, \quad (3)$$

$$A = \frac{1}{2} \begin{bmatrix} 1+\eta_I+\alpha_Q & \alpha_U+i\alpha_V \\ \alpha_U-i\alpha_V & 1+\eta_I-\alpha_Q \end{bmatrix}, \quad (4)$$

and

$$F = \frac{1}{2} \begin{bmatrix} B+(\eta_I+\eta_Q)S & (\eta_U+i\eta_V)S \\ (\eta_U-i\eta_V)S & B+(\eta_I-\eta_Q)S \end{bmatrix}, \quad (5)$$

where the complex quantities α are defined by:

$$\begin{aligned} \alpha_Q &= \eta_Q - i\rho_Q, \\ \alpha_U &= \eta_U - i\rho_U, \\ \alpha_V &= \eta_V - i\rho_V. \end{aligned} \quad (6)$$

Note, that the matrices D and F are hermitian, while A is not. $D(\tau_c)$ describes the polarization state of the light. $A(\tau_c)$ is the Jones matrix (per unit optical depth in the continuum), and describes absorption-, polarization- and birefringent properties of the medium. $F(\tau_c)$ is the source-function matrix. The transport equations (2) can now be written in the form:

$$\frac{dD}{d\tau_c} = A D + D A^\dagger - F, \quad (7)$$

where \dagger denotes a hermitian conjugate (complex conjugation plus transposition).

To solve Eq.(7) I introduce the Jones matrix $T(\tau_c)$ for the entire atmosphere above the level τ_c , which can be found from the differential equation:

$$\frac{dT}{d\tau_c} = A(\tau_c) T(\tau_c), \quad (8)$$

and the boundary condition that $T(0)$ at $\tau_c=0$ is the unit matrix. The hermitian conjugate of T satisfies the equation:

$$\frac{dT^\dagger}{d\tau_c} = T(\tau_c)^\dagger A(\tau_c)^\dagger, \quad (9)$$

I also introduce a matrix $E(\tau_c)$, related to $D(\tau_c)$ by:

$$D = T E T^\dagger. \quad (10)$$

Inserting this into Eq.(7), and using Eqs.(8) and (9), I obtain the following equation for $E(\tau_c)$:

$$\frac{dE}{d\tau_c} = - (T)^{-1} F (T^\dagger)^{-1}. \quad (11)$$

Since $E(\tau_c)$ does not appear on the RHS of Eq.(11), and since the RHS vanishes for large optical depth (due to the exponential increase of T and T^\dagger), the matrix $E(0)$ at $\tau_c = 0$ can be found by direct integration of Eq.(11) over optical depth. However, $E(0) = D(0)$, since $T(0)$ is the unit matrix. The matrix $D(0)$, which describes the Stokes parameters emerging from the atmosphere, can therefore be written as an integral over depth:

$$D(0) = \int_0^\infty (T)^{-1} F (T^\dagger)^{-1} d\tau_c. \quad (12)$$

The argument of this integral,

$$C(\tau_c) \equiv (T)^{-1} F (T^\dagger)^{-1}, \quad (13)$$

is nothing but the contribution function of the matrix $D(0)$. The contribution functions for the emergent Stokes parameters $I(0)$, $Q(0)$, $U(0)$, and $V(0)$ can be found directly from the components of the C-matrix:

$$\begin{aligned} C_I(\tau_c) &= C_{11}(\tau_c) + C_{22}(\tau_c), \\ C_Q(\tau_c) &= C_{11}(\tau_c) - C_{22}(\tau_c), \\ C_U(\tau_c) &= C_{12}(\tau_c) + C_{21}(\tau_c), \end{aligned} \quad (14)$$

$$C_V(\tau_c) = -i[C_{12}(\tau_c) - C_{21}(\tau_c)].$$

These are contribution functions per unit continuum optical depth. To obtain the contribution functions per unit height, Eqs.(14) must be multiplied with κ_c/μ .

To compute the $C(\tau_c)$ -matrix we first have to solve differential equation (8) for $T(\tau_c)$. In general this must be done numerically, e.g. using the Runge-Kutta method. The integration starts at $\tau_c = 0$, and proceeds to larger optical depth. Note, however, that the integration now involves eight coupled equations, since each component of the T-matrix has a real and an imaginary part. In contrast, the traditional method of Eq.(2) involves only four coupled equations. Thus in order to retain information about contribution functions it is necessary to solve a larger system of equations. After the contribution functions of Eq.(14) have been computed, the emergent intensity and polarization can be found by integration of these functions over depth.

It should be pointed out that it is not strictly necessary to use Jones-calculus in order to derive a formal solution of the form (12); a procedure similar to the one above can be applied directly to Eq.(2). However, because of the redundancy in the 4×4 matrices, we then obtain a system of sixteen coupled equations. I conclude that the description in terms of complex 2×2 matrices is more efficient.

3. PHOTOSPHERIC FACULAE

3.1 INTRODUCTION

In the following I present some preliminary results of a study that applies the method of Sect. 2 to a model of photospheric faculae. The small-scale magnetic field of the sun outside sunspots is concentrated into facular elements, or flux tubes, that have typical sizes on the order of a few hundred kilometers, and fieldstrengths of about 1500 gauss (Stenflo, 1973; Harvey, 1977; Zwaan, 1978). Faculae probably consist of vertical flux tubes that are pushed close together by flows in the convective zone below. To first approximation these flux tubes can be considered magnetostatic structures (Spruit, 1976a,b). The magnetic field is contained by a gas-pressure difference between the inside and the outside of the tubes. This pressure difference is a result of the so-called Wilson depression of the levels of constant optical depth inside the tubes, relative to those of the surrounding photosphere. The outward decrease of gas pressure causes the magnetic fieldlines of a flux tube to fan out with height, away from the vertical axis of the tube. Higher up, in the chromosphere and corona, the fields of different faculae merge, and fill the available volume. Waves travelling

upward along flux tubes are believed to play an important role in the heating of the higher atmosphere.

To determine the 3-dimensional structure of photospheric faculae, it is necessary to compare high-resolution observations of these features with predictions based on physically realistic models. The purpose of the present study is to construct a model of an isolated flux tube, and to predict what this structure would look like, when observed with a spectrograph or filtergraph at high spatial resolution.

3.2 FLUX TUBE MODEL

Instead of a more realistic flux-tube model, in which the magnetic field is cylinder-symmetric around a vertical axis, I use here the more simple, 2-dimensional model of a flux sheet, in which the magnetic vector lies in vertical x - z planes, and is independent of the horizontal y -coordinate ($z = -h$ measures depth relative to the level $\tau_{0.5} = 1$; $\tau_{0.5}$ is the continuum optical depth at $\lambda = 0.5 \mu\text{m}$). The flux sheet has boundaries at $x = -R(z)$ and $x = +R(z)$, outside of which the gas is field-free ($R(z)$ is the depth-dependent "radius"). The field inside the boundaries is assumed to be a potential field, so that all electric currents (in the y -direction) are concentrated at the two boundaries. It follows from Ampere's law that the current density per unit length along the boundary is equal to $(4\pi/c)B_0(z)$, where $B_0(z)$ is the fieldstrength just inside the boundary. I assume magneto-static equilibrium, so that the gas pressures $p_i(z)$ and $p_e(z)$ inside and outside the sheet are both independent of x . At the sheet boundary there is a jump in gas pressure, equal to $p_e(z) - p_i(z) = [B_0(z)]^2 / 8\pi$.

To determine the location $R(z)$ of the sheet boundary for given $B_0(z)$, I use the following iterative method. All currents are located at the boundary, so the vector potential $A(x, z)$ of the magnetic field inside the sheet can be expressed as an integral over the boundary surface. For an assumed shape $R_1(z)$ the vector potential $A_1(z)$ on the boundary $x = +R_1(z)$ is computed. The assumed shape is correct only if $A_1(z) = \text{constant}$, so that fieldlines do not cross the boundary. If $A_1(z)$ is not constant, the difference with the $z=0$ level, $\Delta A_1(z) = A_1(z) - A_1(0)$, is used to correct the boundary shape: $R_2(z) = R_1(z) - \Delta A_1(z) / B_0(z)$. Note, that the radius at $z=0$ does not change, so that $R(0)$ is a free parameter of the model.

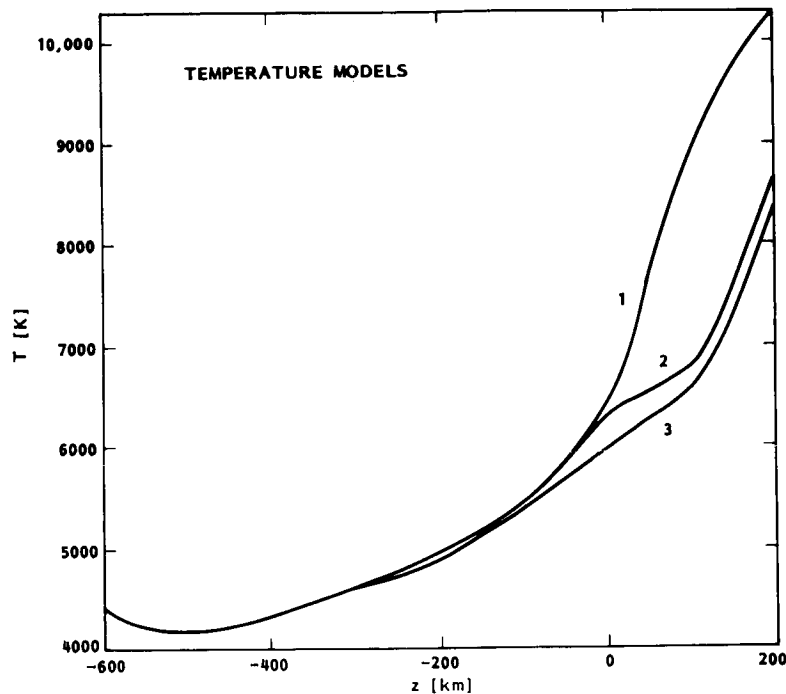


Fig.1 Temperature as function of depth in facular model

The temperature structure of the model is shown in Fig.1. Curve 1 gives the temperature far away from the sheet boundary, adopted from model C of Vernazza et al. (1981) and the convective-zone model of Spruit (1976a). Curve 3 gives the temperature on the mid-plane $x=0$, taken from one of the flux-tube models (model 2) of Spruit (1976b). Spruit's model assumed a tube radius $R(0)$ of 84 km, and a Wilson depression of 100 km, but actually his model was computed for a flux tube with cylindrical geometry. I use this model here only to illustrate the method, and I do not imply that it gives a realistic description of faculae. The gas pressures $p_e(z)$ and $p_i(z)$, necessary to compute the magnetic field, were found by solving the hydrostatic-equilibrium equation for the temperature models of curves 1 and 3. A vertical cross-section of the flux-sheet is shown in Fig.2; the level $\tau_{0.5} = 1$ is indicated by horizontal lines.

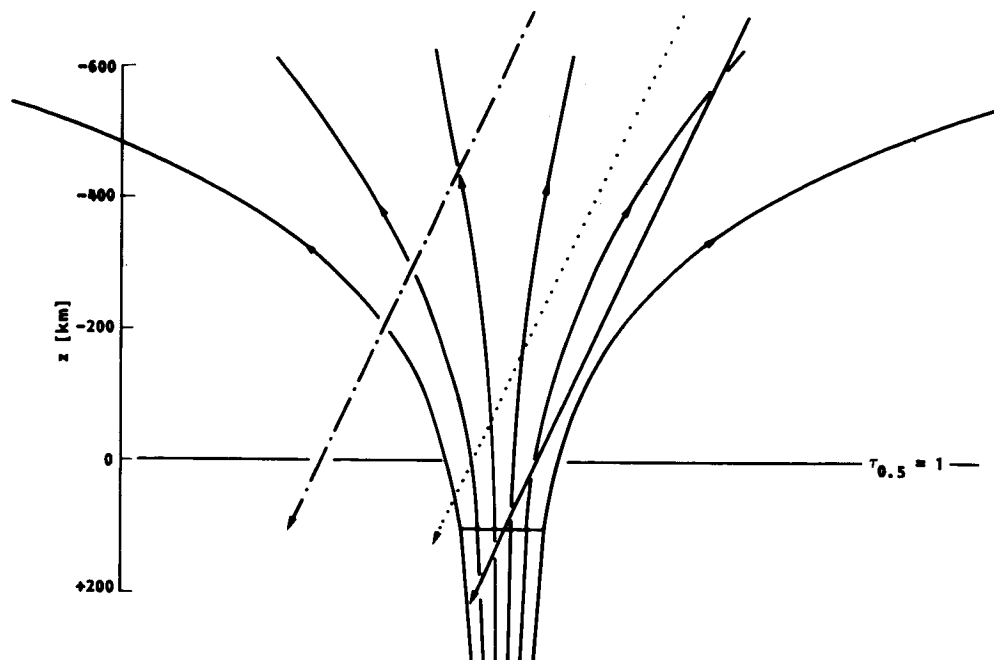


Fig.2 Fieldlines of facular model, and three lines-of-sight

Curve 2 of Fig.1 gives the wall temperature of the gas just outside the sheet boundary (also taken from Spruit, 1976b). This model, and the hydrostatic pressures computed from it, were used for the radiative transport calculations described in the next section. Although the use of two different external models is not consistent with our assumption of hydrostatic balance, the rationale behind the approach is that the external pressure, $p_e(z)$, is determined mostly by the mean stratification of the atmosphere, and not by the immediate surroundings of the flux tube. The temperature reduction in the wall with respect to the mean atmosphere is important, however, for computing realistic intensities of the light emitted by the wall.

3.3 FORMATION OF FeI λ 5250.2

In Fig. 2 three lines-of-sight (LOS), that lie in the x - z plane and make an angle of 25° with the vertical, are also shown. The full-drawn line intersects the $\tau_{0.5}=1$ level in the interior of the flux sheet, while the dotted- and dash-dotted lines cross the left boundary of the sheet at $z=+50$ km and $z=-200$ km. Spectral profiles

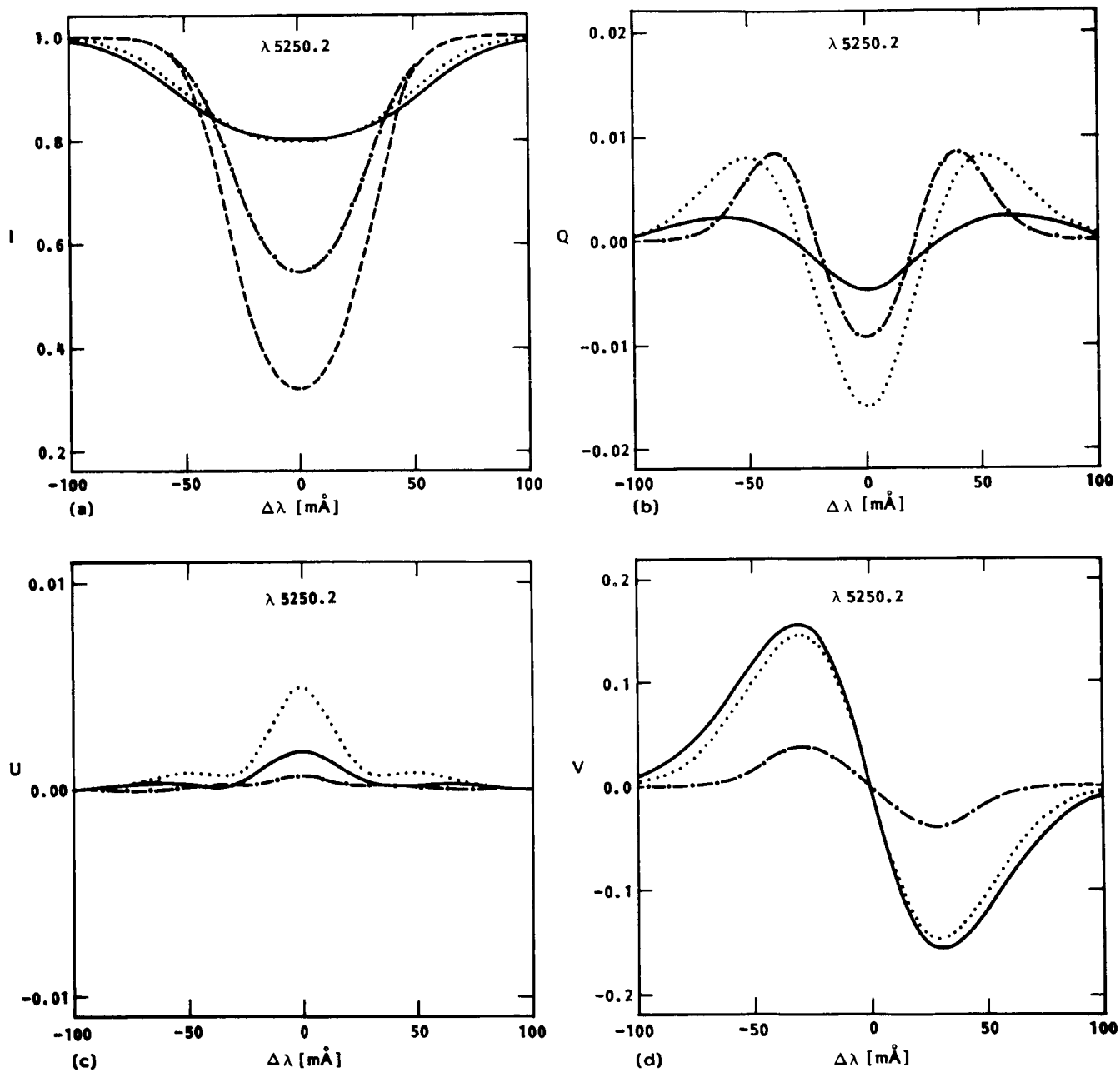


Fig.3 Wavelength profiles of I, Q, U, and V for FeI 5250.2, for three lines-of-sight of Fig.2, and for LOS in quiet sun (dashed curve)

of the line $\text{FeI } \lambda 5250.2$ were computed in LTE for each LOS. The results are shown in Fig. 3, normalized according to the continuum intensity for each LOS. The dashed curve in Fig. 3a is for a LOS in the "quiet" sun away from the flux sheet, although it was still computed with the wall model (curve 2 of Fig. 1). The continuum intensities for the full-, dotted-, and dash-dotted LOS were respectively 1.37, 1.32, and 1.03 times the "quiet" value. Micro-turbulent velocities inside and outside the sheet were taken to be 1 km/s, and macro-turbulence was neglected. Systematic flows were also neglected, so that the wavelength-profiles of I, Q, and U are symmetric, while the profiles of V are anti-symmetric.

Figs. 4a, 4c, and 4e give the contribution functions for the intensity, as measured at line center, in the line wing ($\Delta\lambda = 30 \text{ m\AA}$), and in the continuum, respectively. The curves are normalized to their peak value. The dashed curves again refer to a LOS outside the flux sheet. Figs. 4b and 4d give contribution functions for linear- and circular polarization, measured respectively at line center and in the line wing, where these signals are strongest. In this case the normalization is according to the integral over each curve.

The dashed curve in Fig. 4a shows that the line-center intensity in the quiet sun is formed over a large range in depth, from about $z = -400$ to $z = 0$ km (also see Lites, 1972). For the dash-dotted and dotted LOS I find a large jump in the contribution function at the level where the LOS crosses the wall: because of the lower opacity in the interior, most of the light comes from the external medium behind the wall. The full curves in Figs. 4a,c,e are nearly identical because the spectral line is so weak in the interior of the flux sheet (cf. Fig. 3a).

Figs. 4b and 4d show that the contribution functions for Q and V change sign in the line-forming region. This is due to the fact that Q, U, and V measure intensity differences between two orthogonal modes of polarization (e.g. left- and right circular polarization for V). The opacity in these two modes is generally not equal, so that each mode is formed at a different depth in the atmosphere. Since the modes give opposite contributions to the difference signal, the contribution functions of Q, U, and V consist of a positive and a negative peak that are displaced in depth. The peak at larger depth is larger than the one at smaller depth because the Planck function increases with depth in the photosphere. Thus in LTE the observed polarization is due to the temperature gradient in the atmosphere.

4. DISCUSSION

It was shown that, for the case of polarized light, a formal solution of the transport equations can be obtained. This leads to a natural definition for the contribution functions of Stokes

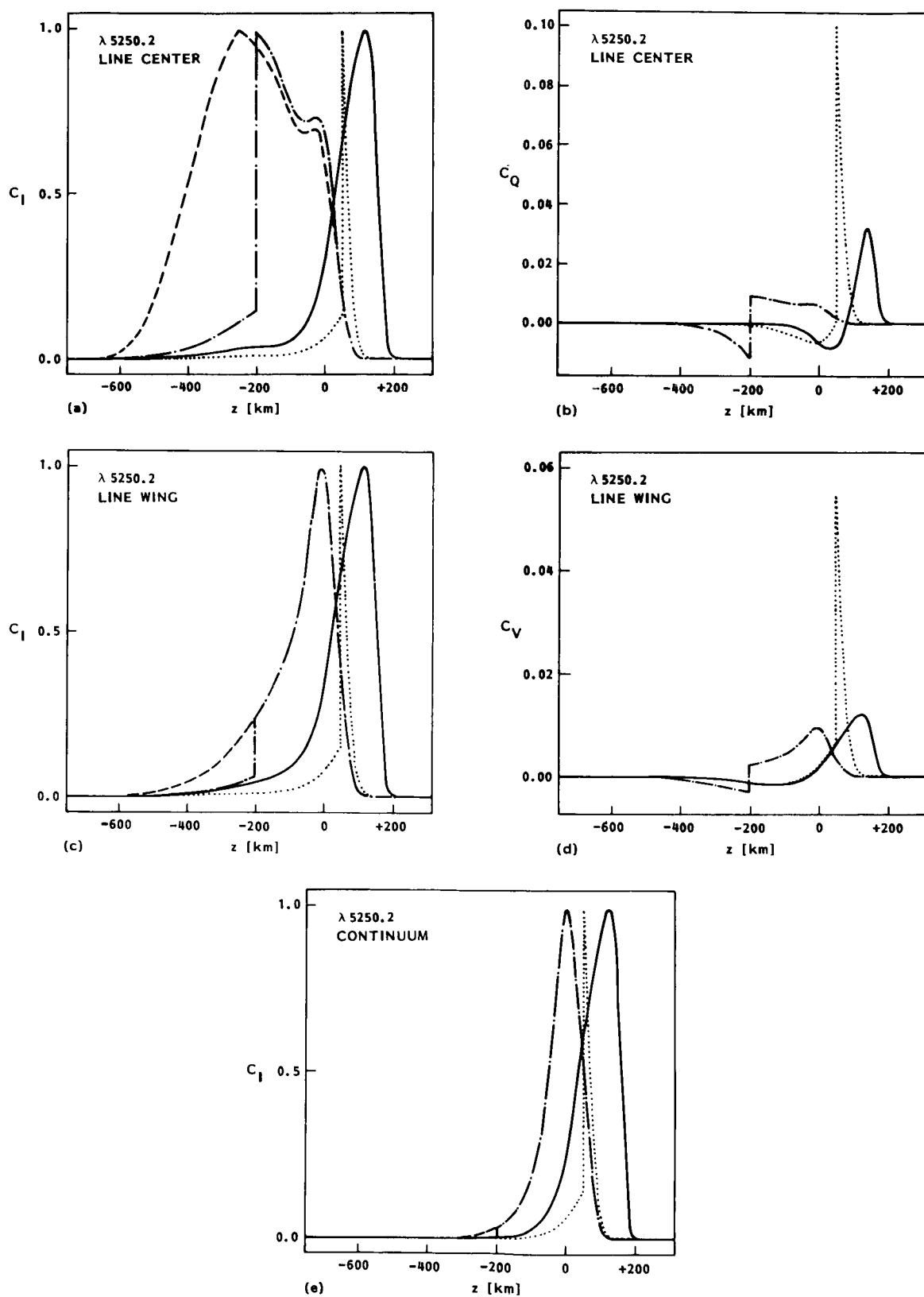


Fig.4 Contribution functions of I, Q, and V for FeI $\lambda 5250.2$

parameters. The present formalism includes non-LTE effects due to scattering, but only in the simplest approximation, in which the scattering is isotropic, with complete redistribution in frequency and polarization. Differences in population between the sublevels with different m quantum numbers in each atomic level (atomic level polarization) have also been neglected. The formalism seems adequate for the description of most non-resonance lines in the photosphere and low chromosphere, in particular the lines of neutral atoms frequently used in magnetic studies.

The facular model indicates that the effects of the walls are important not only for the continuum (Spruit, 1976b), but also for the formation of spectral lines. Thus one-dimensional models (e.g. Unno's equations) are not adequate to interpret the observation. A detailed comparison between models and observations is necessary to determine the thermodynamic structure of faculae.

This work was supported by Lockheed's Independent Research program.

REFERENCES

- Beckers, J.M.: 1969, Solar Phys. 9, 372
- Harvey, J.W.: 1977, in E.A. Mueller (ed.), Highlights of Astronomy, 4, 223
- Kjeldseth-Moe, O.: 1968, Solar Phys. 4, 267
- Landi Degl'Innocenti, E. and Landi Degl'Innocenti, M.: 1972, Solar Phys. 27, 319
- Landi Degl'Innocenti, E. and Landi Degl'Innocenti, M.: 1973, Solar Phys. 31, 299
- Landi Degl'Innocenti, E. and Landi Degl'Innocenti, M.: 1977, Astron. Astrophys., 56, 111
- Landi Degl'Innocenti, E.: 1976, Astron. Astrophys. Suppl. Series 25, 369
- Lites, B.W.: 1972, thesis, University of Colorado and HAO/NCAR
- Spruit, H.C.: 1976a, thesis, University of Utrecht
- Spruit, H.C.: 1976b, Solar Phys., 50, 269
- Staude, J.: 1969, Solar Phys., 8, 264

Stenflo, J.O.: 1971, IAU Symp. 43, "Solar Magnetic Fields", ed. R. Howard, p.101

Stenflo, J.O.: 1973, Solar Phys., 32, 41

Unno, W.: 1956, Publ. Astron. Soc. Japan, 8, 108

Vernazza, J.E., Avrett, E.H., Loeser, R.: 1981, Astrophys. J Suppl. Series, 45, 635

Wittmann, A.: 1974, Solar Phys., 35, 11

Wittmann, A.: 1977, Astron. Astrophys. 54, 175

Zwaan, C.: 1978, Solar Phys., 60, 213

**EFFECT OF POLARIZED RADIATIVE TRANSFER ON THE HANLE MAGNETIC
FIELD DETERMINATION IN PROMINENCES: ANALYSIS OF HYDROGEN
H α LINE OBSERVATIONS AT PIC-DU-MIDI**

V. BOMMIER,* E. LANDI DEGL'INNOCENTI,** J. L. LEROY,*** and S. SAHAL-BRECHOT*

**Observatoire de Paris-Meudon, 92190 Meudon, France*

***Istituto di Astronomia, Universita di Firenze, 50125 Firenze, Italy*

****Observatoire du Pic-du-Midi et de Toulouse, 31400 Toulouse, France*

Abstract. The linear polarization of the Hydrogen H α line of prominences has been computed, taking into account the effect of a magnetic field (Hanle effect), of the radiative transfer in the prominence, and of the depolarization due to collisions with the surrounding electrons and protons. The corresponding formalisms are developed in a forthcoming series of papers. In this paper, the main features of the computation method are summarized. The results of computation have been used for interpretation in term of magnetic field vector measurements from H α polarimetric observations in prominences performed at Pic-du-Midi coronagraph-polarimeter (Leroy, 1981). Simultaneous observations in one optically thin line (He I D₃) and one optically thick line (H α) give an opportunity for solving the ambiguity on the field vector determination.

1. Introduction

The linear polarization of the Hydrogen H α line of solar prominences has been measured in the recent past by Leroy (1981) simultaneously with the He I D₃ polarization, with the aim of obtaining a complete determination of the magnetic field vector by means of the interpretation of the Hanle effect. As a matter of fact, two lines are necessary for the determination of the three components of the field vector, as one line provides only two measured parameters.

Theoretical and computational works have first been devoted to the Hanle effect in the He I D₃ line (Bommier and Sahal-Br  chot, 1978, Bommier, 1980) and of its two components (Landi degl'Innocenti, 1982), and quantitative interpretation of linear polarization measurements of He I D₃ has followed (Leroy et al., 1983, 1984, Athay et al., 1983). After that, the interpretation of the Hydrogen H α measurements has raised a number of new physical problems contrary to He I D₃ which is optically thin: first, the optical thickness of prominences in H α is ~ 1 (Landman and Mongillo, 1979; Engvold, 1978), and the effect of radiative transfer on the linear polarization has to be taken into account. Secondly, owing to the large collisional dipole interaction between the quasi-degenerated levels n, l, j and $n, l\pm 1, j'$ of Hydrogen with the surrounding electrons and protons, the corresponding collisional rates can be large (cf. linear Stark effect in hydrogenic line broadening). The collisional transition probabilities are in fact of the same order of magnitude, for the electronic densities of interest in prominences as the radiative transition probabilities; thus, collisional depolarization has to be taken into account in the interpretation of linear polarization measurements.

In this paper, we summarize the main features of the method which we have used for computing the linear polarization of the H α line of prominences in the presence of a magnetic field. The corresponding formalisms and basic assumptions are developed in a forthcoming series of papers.

II. The Prominence Model

In a recent series of papers, one of us (Landi degl'Innocenti, 1983a,b) has presented a very general formalism on the physics of generation and transfer of polarized radiation in spectral lines. By means of this formalism, it is possible to address a large variety of problems, including the one that is the subject of the present paper, by coupling the radiative transfer equations for polarized radiation to the statistical equilibrium equations for the density matrix elements of a multi-level atomic system. The application of this formalism to the present problem would then require: a) writing the equations of radiative transfer for Hydrogen lines and continua for each direction across the prominence; b) writing the statistical equilibrium equations for the density matrix elements of the Hydrogen atom for each point in the prominence; c) finding the self-consistent solution of the coupled set of equations. This kind of approach, though possible in principle, would however require the solution of a system of equations with such a large number of unknowns that even the fastest computer now available would be inadequate.

In the light of these arguments, it is then necessary, at the present time, to find the solution of the problem through a perturbative method based on the fact that the fractional polarization, as observed in prominences, is always very small ($< 10^{-2}$).

In a first step, which is the zero-order step, the conventional non-LTE problem of Hydrogen-line formation in prominences, is solved, following for instance Heasley and Milkey (1976, 1978). These models have been obtained by solving the coupled equations of radiative transfer and statistical equilibrium for a semi-infinite slab standing vertically on the Sun's surface. The basic assumptions used in these models are the following: the Lyman lines are optically thick across the prominence so that they can be considered in detailed balance; all the other lines are optically thin and in particular $H\alpha$; the Lyman continuum is optically thick, while the subordinate continua are optically thin. However, according to these authors, ~~the non-LTE solutions for slabs and infinite cylinders are virtually identical~~; therefore, as on one hand a real prominence would probably be composed of many such threads separated by hot coronal gas, and on the other hand the observed $H\alpha$ line is not optically thin, it must be assumed that several of these threads are present along the line-of-sight. In fact, Heasley's computations show that the source function in the $H\alpha$ line in one isolated thread has the same value as the mean photospheric intensity at the thread bottom; therefore, we assume that the presence of surrounding threads will not strongly modify the source function in the $H\alpha$ line.

However, if Heasley can state that slabs and infinite cylinders give practically the same results for intensity ratios in line pairs, the same model does not hold for polarization observations. It seems then more appropriate, for the interpretation of polarimetric observations, to retain from Heasley's models only those quantities that are less sensitive to geometrical effects and to model the prominence geometry a-priori. The least parameter-dependent model which can be found is the infinitely sharp model, where the prominence is represented by a collection of threads standing vertically on the solar surface. The slab extends infinitely in the y and z directions (Fig. 1) and is characterized by a finite value τ_0 of the optical depth at the center of $H\alpha$ measured across the prominence. The infinitely sharp assumption means that the optical depth τ_0 is much smaller than the optical depth τ_1 , irrespectively of the height. In other words, the prominence bottom is so small that the bottom of the prominence is always seen as an infinitely sharp line from any height in the prominence.

It results from analytical calculations using a very simplified model of the Hydrogen atom (Bommier and Sahal-Br  chot, to be published), and from numerical tests, that the linear polarization of the $H\alpha$ line depends only on the ratio $S(\text{Ly}\beta)/S(\text{Ly}\alpha)$ of the source functions in the prominence. The $H\alpha$ polarization depends obviously also on the source function $S(H\alpha)$ and on the anisotropy of the radiation in this line. In the Heasley's series of models, $S(\text{Ly}\beta)/S(\text{Ly}\alpha)$ and $S(H\alpha)$ are in fact nearly independent of the geo-

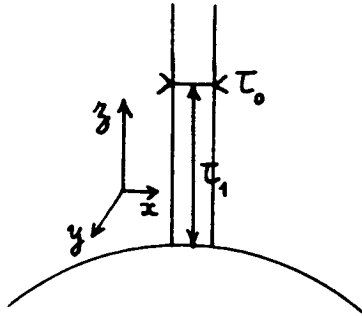


Fig. 1

and e_1 and e_2 are two unit vectors perpendicular to the direction Ω . λ is the wavelength of one of the nearly optically thin lines ($H\alpha, H\beta, P\alpha$). For those directions that are not coming from the solar photosphere, one has:

$$I(\lambda, \vec{\Omega}) = S_0(\lambda) [1 - e^{-\tau_\lambda}] ;$$

for the other directions:

$$I(\lambda, \vec{\Omega}) = S_0(\lambda) [1 - e^{-\tau_\lambda}] + I_0(\lambda, \vec{\Omega}) e^{-\tau_\lambda} ,$$

where $S_0(\lambda)$ is the source function and $I_0(\lambda, \vec{\Omega})$ is the photospheric intensity in the $\vec{\Omega}$ direction. As for the optically thick Lyman lines, the radiation field in the prominence is isotropic and the tensor is a diagonal matrix.

III. Computation of the $H\alpha$ Linear Polarization

The next step is to solve the statistical equilibrium system of equations for the atomic density matrix (Bommier and Sahal-Br  chot, 1978, Bommier, 1980, Landi degl'Innocenti, 1982). The radiation field tensor defined in the preceding section is nothing more than the incident photon density matrix which has to be coupled to the atomic density matrix, in the presence of a magnetic field.

The atomic density matrix of the Hydrogen atom has been expanded over a basis which diagonalizes the fine-structure and the magnetic field interaction. The validity of the computation is then extended up to field strengths of 100 Gauss. The hyperfine structure also has been taken into account in a preliminary calculation, but turns out to be negligible with respect to the measurement inaccuracies.

The statistical equilibrium results from radiative transitions and also from transitions between fine-structure levels n, l, j and $n, l \pm 1, j'$ due to collisions with the surrounding electrons and protons. The corresponding collisional rates are of the same order of magnitude as the radiative rates, for typical electronic densities of prominences, owing to the large dipole interaction between the quasi-degenerated levels in Hydrogen. We have computed the collisional transition probabilities with a semi-classical perturbation method, in the impact approximation which is valid at typical electron densities in prominences.

metrical parameters of the model. Therefore, using their values, the tensor components of the radiation field have been computed in each point of a grid of points in the prominence:

$$\phi_{ij}(\lambda) = \oint \frac{d\vec{\Omega}}{4\pi} I(\lambda, \vec{\Omega}) E_{ij}(\vec{\Omega}) ; (i,j) = (x,y,z) ;$$

where $E_{ij}(\vec{\Omega})$ is a 3×3 geometrical tensor defined by:

$$E_{ij}(\vec{\Omega}) = \frac{1}{2} [(\vec{e}_1)_i (\vec{e}_1)_j + (\vec{e}_2)_i (\vec{e}_2)_j] ,$$

The effect of collisions with protons has been found to be 10 times larger than the effect of collisions with electrons. Owing to the isotropic distribution of electrons and protons, the effect of collisions is to decrease the anisotropy in the Zeeman sublevels which is responsible for the polarization of the emitted radiation.

The statistical equilibrium equations have been solved for the levels $n=1$ to $n=3$ for computations of the $H\alpha$ polarization. The effect of radiative cascades from upper levels, which is important for line intensity computations as shown by Heasley and Milkey (1978), has been found not to be important for polarization computations, and can be neglected.

The computation gives the elements of the atomic density matrix for each point of the grid. The 4×4 absorption matrix and the emission vector, which enter the radiative transfer equations in matrix form (Landi degl'Innocenti, 1983a,b), are derived from the atomic density matrix. The transfer equation for polarized radiation is then numerically solved and the $H\alpha$ emitted polarization is obtained.

IV. Results of Computation

Figure 2 displays the integrated intensity emitted in $H\alpha$ as a function of the optical thickness along the line-of-sight τ_L computed from our model. The agreement between the theoretical curve and the observational results of Landman and Mongillo (1979) and Engvold (1978), reflects, as expected, the differences already discussed by both observers, between these observational results and the results from Heasley's series of models.

Figure 3 shows the degree of linear polarization in zero magnetic field as a function of the integrated intensity $E(H\alpha)$, for various angles θ_L between the prominence long axis and the line-of-sight.

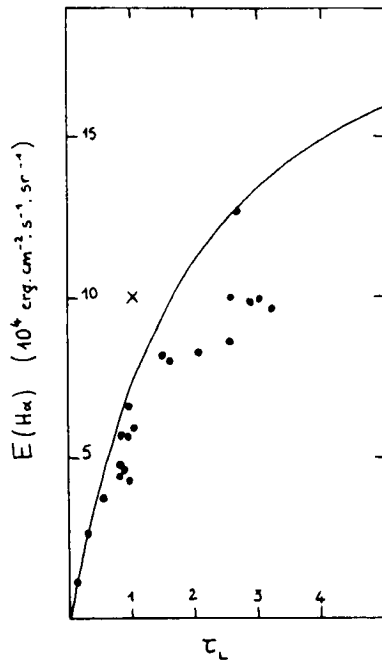


Fig. 2. Integrated intensity emitted in $H\alpha$ as a function of optical thickness along the line-of-sight. ●: empirical values deduced by the interpretation of line-profile measurements (Landman and Mongillo, 1979); x: value estimated from line intensity ratios measurements.

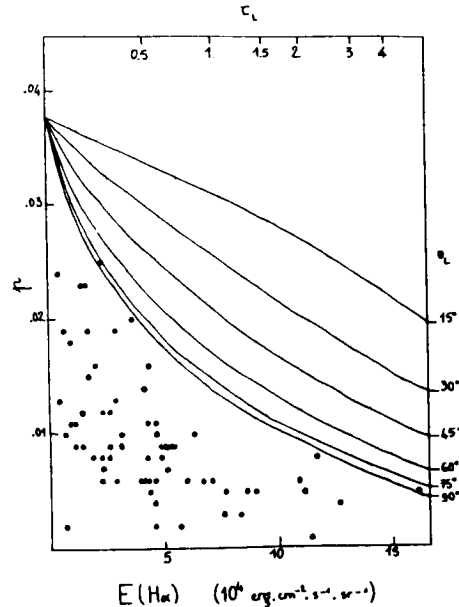


Fig. 3. Linear polarization degree in zero magnetic field as a function of the emitted intensity in $H\alpha$, computed for a height of 60 arc sec above the solar limb and for various angles θ_L between the line-of-sight and the prominence long axis. ●: $H\alpha$ linear polarization measurements in the same altitude range (Leroy, 1981).

The computations have been done for a typical height of 60 arc sec above the limb. The depolarization due to increasing optical thickness along the line-of-sight is clearly shown, in agreement with the behaviour of observational data (Leroy, 1981, Fig. 5 of that paper); it should be reminded, however, that, in the observed polarization, the depolarizing effect of the magnetic field is also involved.

In the presence of a magnetic field, the degree of polarization is decreased and the polarization direction is rotated from the polarization direction in zero magnetic field. These are the main features of the Hanle effect. This can be represented in diagrams where the abscissa is the rotation of the polarization direction and the degree of polarization, normalized to the polarization degree in zero magnetic field, is the ordinate. The diagrams given on Fig. 4 have been computed for an optical thickness in $H\alpha$ $\tau_L = 2$ along the line-of-sight, and an angle $\theta_L = 30^\circ$ between the line-of-sight and the prominence long axis. The magnetic field is assumed to be horizontal. The field azimuth is referred with respect to the line-of-sight and two opposite values of the azimuth angle correspond to two symmetrical field vectors with respect to the line-of-sight. The two diagrams given on Fig. 4 correspond respectively to positive and negative azimuth angles of the magnetic field. By comparing these two diagrams, it can be seen that two field vectors symmetrical with respect to the line-of-sight have not the same effect on the $H\alpha$ linear polarization. That is an optical thickness effect; in the optically thin case, the symmetry in 180° rotation around the line-of-sight implies that two symmetrical field vectors have the same effect on the linear polarization.

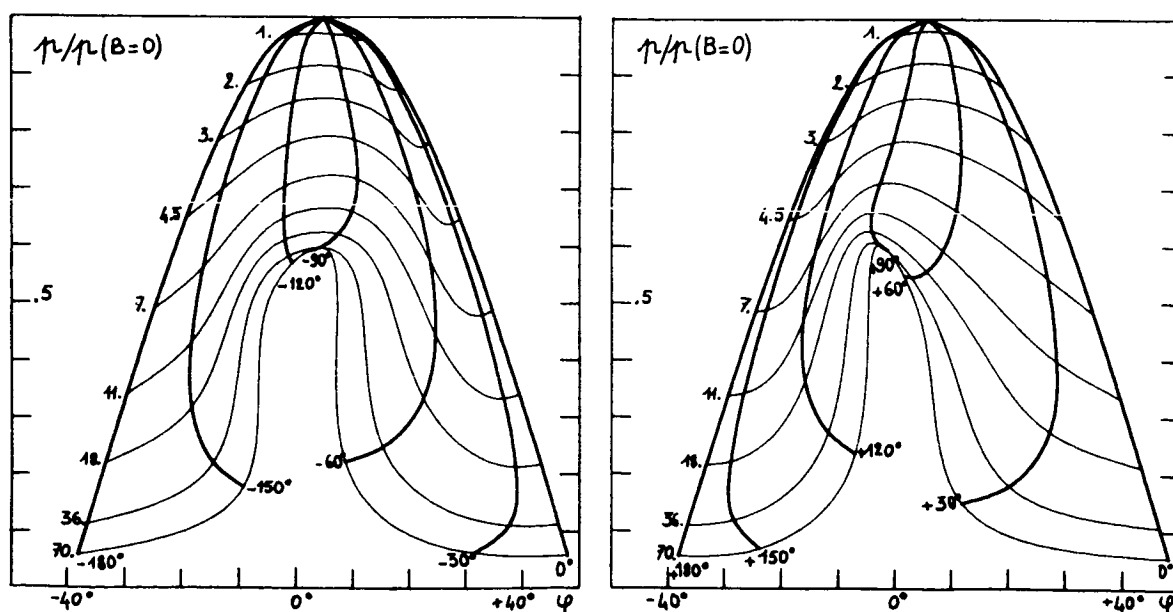


Fig. 4. Hanle effect in prominence $H\alpha$ emission, computed at height of 60 arc sec. above the limb. In abscissae is the rotation φ of the polarization direction from the limb direction. In ordinate is the polarization degree, normalized to the zero-field polarization degree .02. The bold curves correspond to constant azimuth of the field vector with respect to the line-of-sight. The light curves correspond to constant field strength, in Gauss. The magnetic field and the line-of-sight lie in the local horizontal plane. The optical thickness of the prominence along the line-of-sight is 2.; the angle between the line-of-sight and the prominence long axis is 30° .

V. Analysis of H α Observations at Pic-du-Midi: Solution of the 180° Ambiguity

As a consequence of the preceding section, simultaneous observations in one optically thin line and one optically thick line can be used to solve the 180° ambiguity in the field vector determination. Such observations have been performed with the Pic-du-Midi coronagraph-polarimeter on 20 prominences of high altitude. The polarization of the He I D₃ λ 5876 line, which is optically thin, and the polarization of the H α line have been measured quasi-simultaneously. The He I D₃ observations have been interpreted by using the polarization computations of Bommier and Sahal-Br  chot (1978) and Bommier (1980). The magnetic field has been assumed to be horizontal. Two field vectors have been provided and the H α polarization has been computed for each of the two field vectors, following the method described in this paper. Then, the computed H α polarization has been compared with the observed one. The angle between the line-of-sight and the prominence long axis has been derived from identification of the prominence with a filament observed earlier or later on the disk. In more than 50% of cases, the difference between the observed and computed polarization has been found to be significant for one of the two field vectors, with respect to the measurements errors. Accordingly, the ambiguity can be solved for those prominences. The orientation of the remaining field vector with respect to the polarity of the neighbouring photospheric field has been found to be consistent with a magnetic structure of the Kuperus and Raadu type, i.e. the magnetic field crosses the prominence from the negative photospheric polarity towards the positive photospheric polarity.

This result is in agreement with the results that we have obtained with another method for solving the ambiguity, which is based on statistical analysis of a large sample of prominences (Bommier, Leroy, Sahal-Br  chot, this workshop; Leroy et al., 1984). The prominences at high altitude have been statistically found to be of the Kuperus and Raadu type.

References

- Athay, R. G., Querfeld, C. W., Smartt, R. N., Landi degl'Innocenti, E., Bommier, V.; 1983, *Solar Phys.*, **89**, 3.
- Bommier, V., Sahal-Br  chot, S.; 1978, *Astron. Astrophys.*, **69**, 57.
- Bommier, V., Sahal-Br  chot, to be published.
- Bommier, V.; 1980, *Astron. Astrophys.*, **69**, 57.
- Engvold, O.; 1978, *Solar Phys.*, **71**, 285.
- Heasley, J. N., Milkey, R. W.; 1976, *Astrophys. J.*, **210**, 827.
- Heasley, J. N., Milkey, R. W.; 1978, *Astrophys. J.*, **221**, 677.
- Landi degl'Innocenti, E.; 1982, *Solar Phys.*, **79**, 291.
- Landi degl'Innocenti, E.; 1983a, *Solar Phys.*, **85**, 3.
- Landi degl'Innocenti, E.; 1983b, *Solar Phys.*, **85**, 33.
- Landman, M., Mongillo, D.; 1979, *Astrophys. J.*, **230**, 581.
- Leroy, J. L.; 1981, *Solar Phys.*, **71**, 285.
- Leroy, J. L., Bommier, V., Sahal-Br  chot, S.; 1983, *Solar Phys.*, **83**, 135.
- Leroy, J. L., Bommier, V., Sahal-Br  chot, S.; 1984, *Astron. Astrophys.*, **131**, 33.

The Effect of Line Damping, Magneto-optics and Parasitic Light
on the Derivation of Sunspot Vector Magnetic Fields

A. SKUMANICH/High Altitude Observatory - NCAR¹

B.W. LITES/Sacramento Peak National Solar Observatory²

The least square fitting of Stokes observations of sunspots using a Milne-Eddington-Unno model appears to lead, in many circumstances, to various inconsistencies such as anomalously large doppler widths and, hence, small magnetic fields which are significantly below those inferred solely from the Zeeman splitting in the intensity profile. We find that the introduction of additional physics into the model such as the inclusion of damping wings and magneto-optic birefringence significantly improves the fit to the Stokes parameters. Model fits excluding the intensity profile, i.e. of both magnitude as well as spectral shape of the polarization parameters alone, suggest that parasitic light in the intensity profile may also be a source of inconsistencies. We discuss the consequences of the physical changes on the vector properties of the field derived from the Fe I $\lambda 6173$ line for the 17 November 1975 spot as well as on the thermodynamic state. We find a doppler width $\Delta\lambda_D = 25\text{m\AA}$, consistent with a low spot temperature and micro-turbulence, and a damping constant of $a = 0.2$. Magneto-optics prove to be important whenever the scaling law, $(\Delta\lambda_H/\Delta\lambda_D)\eta_0 \geq 3$, is satisfied. Here $\Delta\lambda_H$ is the Zeeman splitting, and η_0 , the relative optical thickness (to the continuum) at line center.

¹ The National Center for Atmospheric Research is sponsored by the National Science Foundation.

² Operated by AURA, Inc., under contract with the National Science Foundation.

THE INFERENCE OF VECTOR MAGNETIC FIELDS FROM POLARIZATION
MEASUREMENTS WITH LIMITED SPECTRAL RESOLUTION

Bruce W. Lites
National Solar Observatory
National Optical Astronomy Observatories¹
Sunspot, New Mexico 88349 USA

and

Andrew Skumanich
High Altitude Observatory
National Center for Atmospheric Research²
P.O. Box 3000
Boulder, Colorado 80307 USA

ABSTRACT

We present a method for recovery of the vector magnetic field and thermodynamic parameters from polarization measurements of photospheric line profiles measured with filtergraphs. The method is based on the simplifying assumptions of Unno (1956) and Rashkovsky (1962, 1967), but includes magneto-optic effects and may be utilized on data sampled at arbitrary wavelengths within the line profile. We explore the accuracy of this method through inversion of synthetic Stokes profiles subjected to varying levels of random noise, instrumental wavelength resolution, and line profile sampling. We also investigate the level of error introduced by the systematic effect of: 1) profile sampling over a finite fraction of the 5-minute oscillation cycle, and 2) inaccuracy in the knowledge of the instrumental wavelength resolution. The results presented here are intended to guide instrumental design and observational procedure. Below, we summarize some of the results:

- 1) One may recover the magnetic field strength and direction to good accuracy (5% or better) using filters with good to excellent spectral purity ($\text{FWHM} < 100 \text{ m\AA}$) under the following conditions:
 - a) strong fields ($|B| > 1000 \text{ Gauss}$),
 - b) good coverage of the line profile, i.e. at least two sampling points per instrumental resolution width,
 - c) a signal-to-noise ratio in line with that expected from modern panoramic detectors.

¹Operated by the Association of Universities for Research in Astronomy, Inc. under contract AST 78-17292 with the National Science Foundation.

²NCAR is sponsored by the National Science Foundation.

Tradeoffs are possible in all of the above conditions in order to arrive at any level of accuracy.

2) The thermodynamic parameters derived from the fit are generally much more poorly determined than the magnetic field parameters.

3) Reasonable uncertainty may be tolerated in the knowledge of the filter wavelength response with little influence on the derived magnetic fields.

4) Sampling the Stokes profiles during a finite fraction of the solar 5-minute oscillation cycle introduces a small but tolerable error in the magnetic field parameters.

We also discuss the stability of this method against false solutions and failure to converge.

I. INTRODUCTION

Modern solar physics has an unquestioned need for reliable measurements of vector magnetic fields at a number of heights in various solar features. Such measurements could illuminate problems concerning, for example, electric currents in active regions, field geometry of sunspots, magnetic energy dissipation in flares, and the structure of network magnetic fields with height, to name only a few. Measurements of the full Stokes polarization parameters within solar line profiles are difficult to obtain, and once obtained they are also difficult to interpret. Significant progress has recently been achieved with the first-order interpretation of photospheric sunspot magnetic fields from the High Altitude Observatory/Sacramento Peak Observatory Stokes polarimeter (Skumanich and Lites 1984). That work indicates that high spectral resolution Stokes profiles of spectral lines may be used to successfully extract magnetic field strength and orientation using an Unno (1956)/Rachkovsky (1962,1967) approach modified to include magneto-optical rotation of the polarization. The primary drawback of that type of data seems to be the limited spatial and temporal resolution of the profiles. Modern tunable filters offer the possibility of improving both the spatial and temporal resolution of the polarization measurements, but they do so only at the expense of the spectral resolution. Such a filter vector magnetograph is now in operation (Haygard et al. 1982), albeit at a site where truly high spatial resolution may only rarely be obtained. As several groups are contemplating how to approach the design of a next-generation Stokes facility, and one such group is in fact entering a detailed engineering phase (Rayrole 1982), it is now appropriate to understand quantitatively how the tradeoff between spectral resolution and inferred magnetic field may be met realistically given the noise limitations of modern electronic detectors.

Here we first show how an inversion program (Skumanich and Lites 1984: a modified and improved version of the method of Auer, Heasley, and House 1977) may be modified in a completely consistent way to analyze line polarization measurements taken with arbitrary spectral resolution. We then present results of nu-

merical experiments designed to assess the accuracy to which the magnetic field and thermodynamic parameters may be extracted from the synthetic line polarization profiles, when subjected to varying levels of spectral resolution and noise. Next, through similar numerical simulations of observations we estimate the inaccuracies of inferred magnetic field strengths and orientations introduced by systematic effects of the observation process: inaccurate knowledge of the instrumental resolution function in wavelength and profile distortion introduced by sampling the line profile over a significant fraction of the 5-minute oscillation cycle. Finally, we present a simulation of sunspot magnetic field measurements at several disk positions to test the ability of our method to avoid false solutions for the magnetic and thermodynamic parameters. We defer discussion of the influence of nonuniformity in the model atmosphere a separate study (Skumanich, Rees, and Lites 1984). In that paper we address the question of the influence of gradients in velocity, in the ratio of line to continuum opacity, in the magnetic field and direction, and in the slope of the source function.

II. A METHOD FOR ANALYZING STOKES PROFILES WITH LIMITED SPECTRAL RESOLUTION

The method of inversion of coarse analysis of Stokes profile data employed by Skumanich and Lites (1984) works well for the high resolution spectra described therein, but when these spectra are degraded, even by a bandpass as narrow as admitted by the best modern filters, and then subjected to the same analysis, serious errors in the derived magnetic fields may be expected. We compensate for any arbitrary instrumental resolution function by building this function into the analysis from the beginning. An alternate approach might be to restore the spectral line profiles by Fourier methods (Brault and White 1971), but these methods have the drawback that the restoration is compromised greatly by instrumental noise of the magnitude expected from modern panoramic detectors. In addition, Fourier restoration of the data requires that the Stokes profiles be sampled rather frequently in wavelength and well outside the line profile itself. This restriction would hinder the rapid scanning of a line profile with a non-uniform and coarse wavelength sampling possible with modern tunable filters. The method of analysis described below has none of these disadvantages.

Lites and Skumanich (1984) have described some details of the non-linear least-squares routine used by Skumanich and Lites (1984) to derive the magnetic field and thermodynamic parameters from a set of observed Stokes profiles. They give a description of the set of derivatives of the theoretical Stokes profiles with respect to the model parameters which are used in the least-squares algorithm. Our method convolves the wavelength profiles of these derivatives, as well as the theoretical Stokes profiles, with the known wavelength resolution function. Since the theoretical profiles and derivatives may be calculated to arbitrary precision on a fine wavelength grid, they may be convolved with any instrumental resolution function using standard Fourier techniques. The least-squares fit may then proceed using degraded theoretical Stokes profiles and derivatives, plus the observed profiles sampled with an arbitrary wavelength spacing.

If we consider the possibility that the filter may be tuned through the line profile with a rather coarse spacing (i.e. so that the profile is critically

sampled with respect to the bandpass of the filter), the earlier technique of determining the line center wavelength and symmetrizing the profiles prior to the least-squares routine is no longer tenable. We have therefore added another parameter to the least-squares fitting routine: the position of the center of the line. The fit is now carried out over the entire line profile, not just the symmetrized half-profile as was done by Skumanich and Lites (1984). The stability of least-squares fitting routines usually decreases with the number of parameters to be fit, but our computations show that the line center position is certainly the most stable and well-determined of the 9 possible parameters. Moreover, a good initializing guess for this parameter may be had readily from the data. We have found that the addition of this parameter enhances rather than degrades the quality of the fit since the algorithm previously used to find line center was somewhat less sophisticated than necessary.

Tests of this method on synthetic noise-free data show that the fit parameters may be recovered by the least-squares fitting routine to machine accuracy, even when the initial guess for the parameters is far away from the solution, and even when the Stokes profiles are severely smeared and coarsely sampled. The least-squares routine is therefore quite stable and rapidly convergent for this application.

III. ACCURACY OF THE METHOD WITH NOISY STOKES POLARIZATION MEASUREMENTS

The way in which we add random noise to the synthetic Stokes profiles is governed by the way in which the measurements would be taken. We presume that images are recorded in a filtergraph mode with a characteristic noise level N . Furthermore, we presume that several images are recorded in this mode in various states of linear and circular polarization, so that typically two images are added to get the I (intensity) Stokes parameter, and two images are subtracted to obtain the Q , U , and V Stokes parameters. The noise level in I will then be $N\sqrt{2}$, and in Q , U , and V it will be $2N$. We identify the rms noise relative to the continuum level I_c outside the line in one of the polarization states as σ . We then assume that the rms noise level at any point within the line profile scales as the intensity parameter I at that wavelength: $N = \sigma I_c (I/I_c)^{1/2}$. Noise fluctuations are introduced into the profiles by adding random numbers with a correctly normalized Gaussian distribution about a zero mean. The noise imposed upon all of the Stokes profiles using this scaling may differ slightly from that present in actual measurements of line polarization in instances where the line profile is highly split by the magnetic field. However, we believe that this scaling carries the essential features of the wavelength dependence of the noise. Figure 1 shows the four Stokes profiles for Fe $\lambda 6173$ in the standard strong field case (see below). Also shown are these profiles convolved with three Gaussian functions of widths 30, 75, and 120 mÅ (or full-width at half-maximum, FWHM, of 50, 125, and 200 mÅ, respectively). Superimposed on the smeared profiles are profiles subjected to a noise level of $\sigma = 0.003$: a typical value for modern electronic panoramic detectors. We also presume that systematic errors in the measurement process (due to improper image flat-field correction or improper calibration of the linear polarization, as examples) are insignificant relative to the random measurement noise, aside from those systematic errors considered in Section IV.

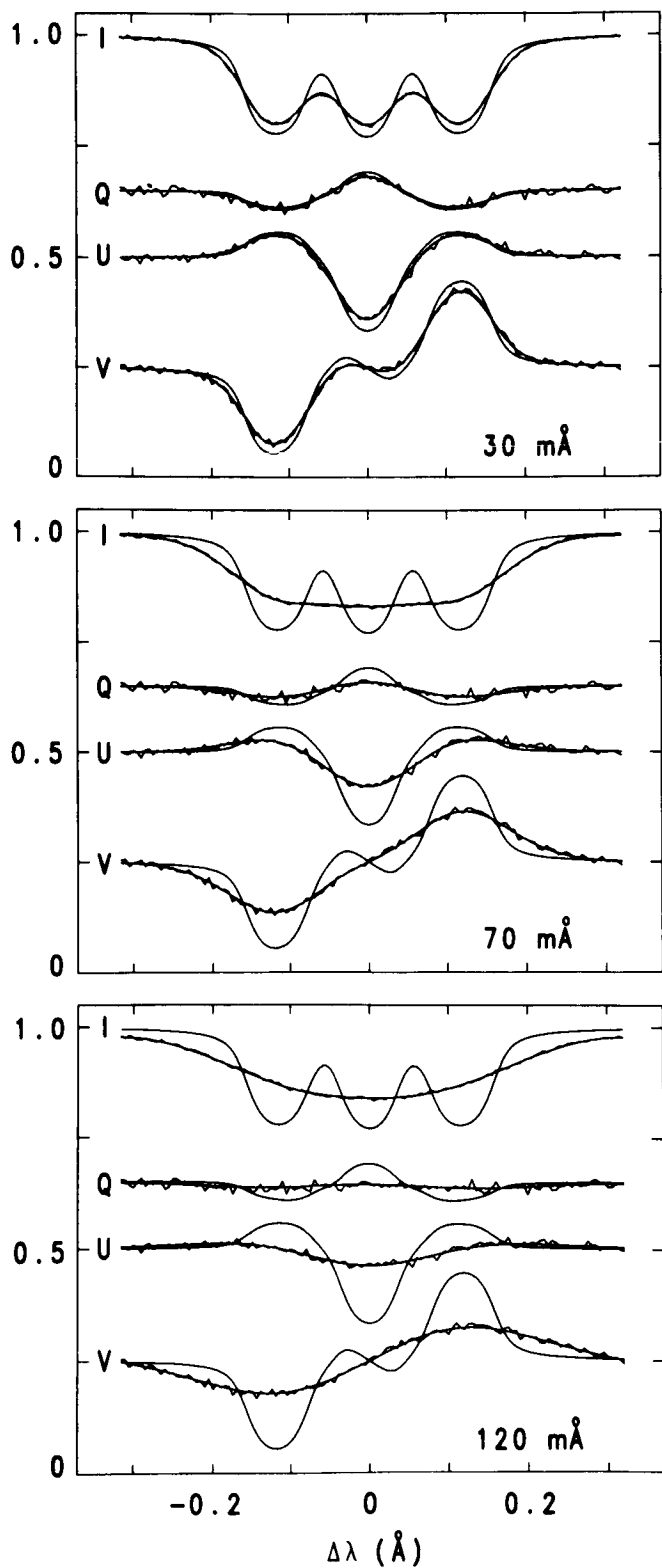


FIGURE 1. Simulated Stokes profiles are shown for the Fe I line $\lambda 6173$. Superimposed are the theoretical profiles for the standard strong field case profiles smeared with a Gaussian function of the width shown in each panel (the FWHM is 1.665 times the value shown), and the profiles with a noise level $\sigma = 0.003$ added. The intensity scale for I is shown relative to the continuum; the Q, U, and -V profiles are displaced from zero by arbitrary amounts. Note: the V Stokes profiles shown here are actually -V, since that is the sign of the real data we have used to date with this data reduction program. The standard case has $|\mathbf{B}| = 2500 \text{ G}$, $\phi = 45^\circ$, $\phi = 60^\circ$, $\eta = 10$, $\mu_B = 10^6$, $\Delta\lambda_D = 25 \text{ mÅ}$, $a = 0.2$.

In our numerical study of the accuracy of extracting magnetic field values from synthetic data we are dealing with a great number of free parameters. There are 8 parameters that define the Stokes profiles in our representation of the Unno (1956) formulation (line center position, the ratio of line center to continuum opacity η_0 , the slope of the source function with optical depth in the line μB_1 , the Doppler width $\Delta\lambda_D$, the damping width a , the absolute value of the magnetic field $|B|$, the field inclination with respect to the line of sight ψ , and the field azimuth ϕ). We also have the level of noise introduced σ , the width of the instrumental resolution function (denoted hereafter as the smearing width), and the sampling interval of the profiles. For each trial with a given combination with these parameters we have chosen to make 20 realizations of the noise in order to get a meaningful value for the rms error and bias of the fit. In order to reduce the computing effort, we have chosen to initialize the non-linear least squares fit with the correct values of the magnetic field and thermodynamic parameters, and with a line center position 20 mÅ away from the true line center in all cases except where noted in the following. The results we have obtained are thus the best one can do, since initializing far from the solution might, in isolated cases, result in a failure to converge to the correct solution in a finite number of iterations, or it might result in convergence to a false solution.

In Figure 2 we show the influence of noise and smearing width on the rms error in the derived magnetic field strength and direction for the standard strong field case ($|B|=2500$ Gauss, inclination angle $\psi = 45^\circ$, field azimuth $\phi = 60^\circ$, $\eta_0 = 10$, $\mu B_1 = 10^6$, $\Delta\lambda_D = 25$ mÅ, $a = 0.2$). These parameters are not unreasonable for the magnetic line Fe I $\lambda 6173$ in a sunspot umbra (Skumanich and Lites 1984). Each curve in Figure 2 gives the rms error (in Gauss for $|B|$, or in degrees for the two angles) as a function of the rms noise level introduced into the Stokes profiles by the method described above. The horizontal lines denote the 1% and 5% levels. Unless otherwise noted, the smearing width given in mÅ indicates the 1/e half-width of the Gaussian function. To convert this smearing width to the FWHM, one must therefore multiply these widths by 1.665. The line profiles were assumed to be sampled every 5 mÅ for the analysis resulting in Figure 2. This analysis indicates several important points: 1) One may expect to recover the magnetic field parameters in the strong field case to high accuracy (1% or better) with modern high-resolution filters (FWHM i.e. 125 mÅ) using only single exposures of a CCD detector. 2) The accuracy of the measurement is very sensitive to the noise level, and somewhat less sensitive to the smearing width. Curves similar to those of Figure 2 for the thermodynamic parameters η_0 , μB_1 , $\Delta\lambda_D$, and a exhibit a very similar appearance, but the thermodynamic parameters are in general much more poorly determined than the magnetic field parameters. For example, with $\sigma = 0.003$, a 5% error is already present in η_0 and the damping a , and 1% in μB_1 and $\Delta\lambda_D$ in the non-smeared case.

Instrumental or observational constraints may prohibit sampling of the line profile on as fine a grid as adopted for the calculations leading to Figure 2. In Figure 3 we illustrate the influence of profile sampling on the accuracy of the inferred magnetic field for the standard case. Here we have adopted an rms noise level of $\sigma = 0.003$, and each curve represents the rms error as a function

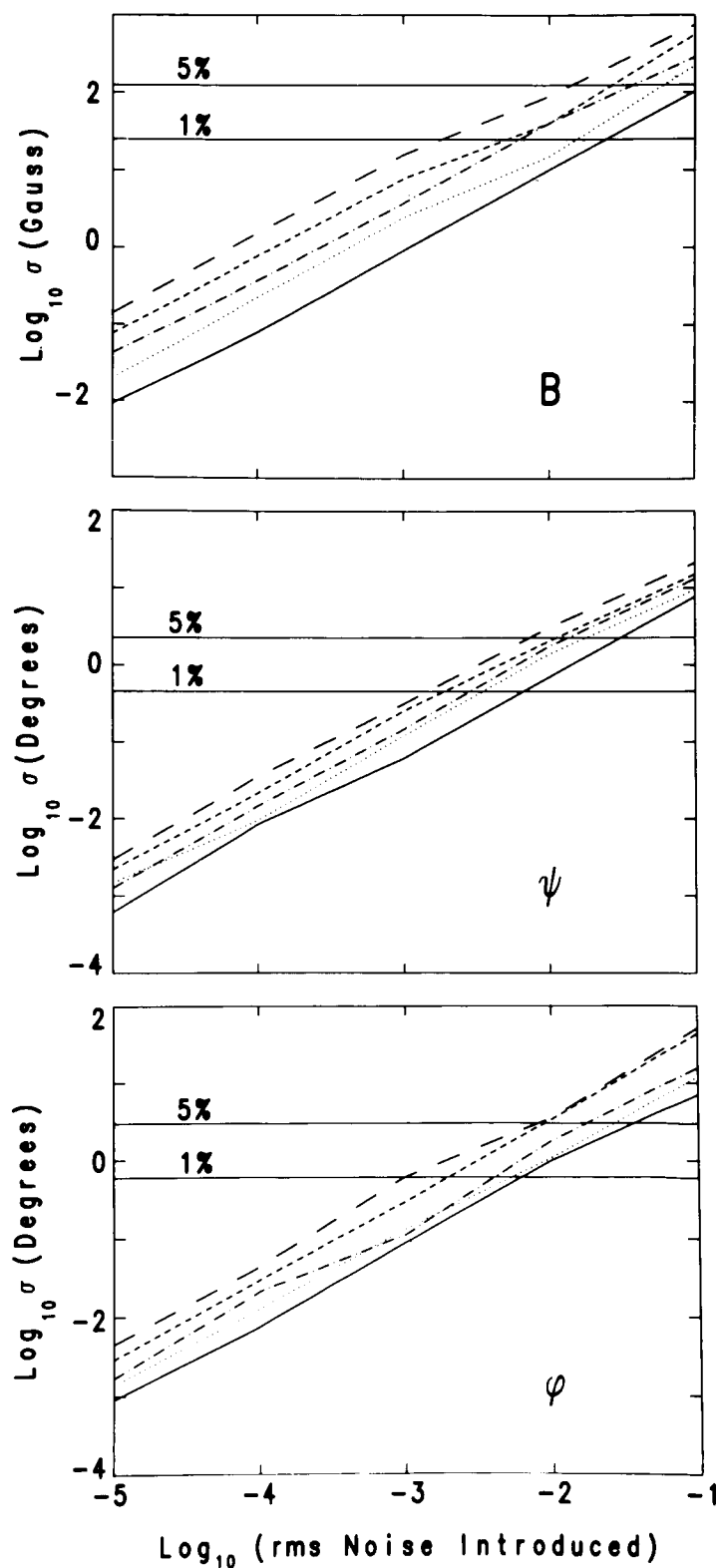


FIGURE 2. The rms error expected in the magnetic field strength (top panel), the angle of inclination with respect to the line-of-sight ψ (middle panel), and the azimuthal angle ϕ (bottom panel) are shown as a function of the noise level σ introduced into the profiles. The solid, dotted, short-dashed, short-dashed, and long-dashed lines represent rms errors for Gaussian filter widths of 0, 50, 70, 100, and 140 mA, respectively. The profiles were all sampled at 5 mA intervals. The horizontal lines show the 1% and 5% rms error levels.

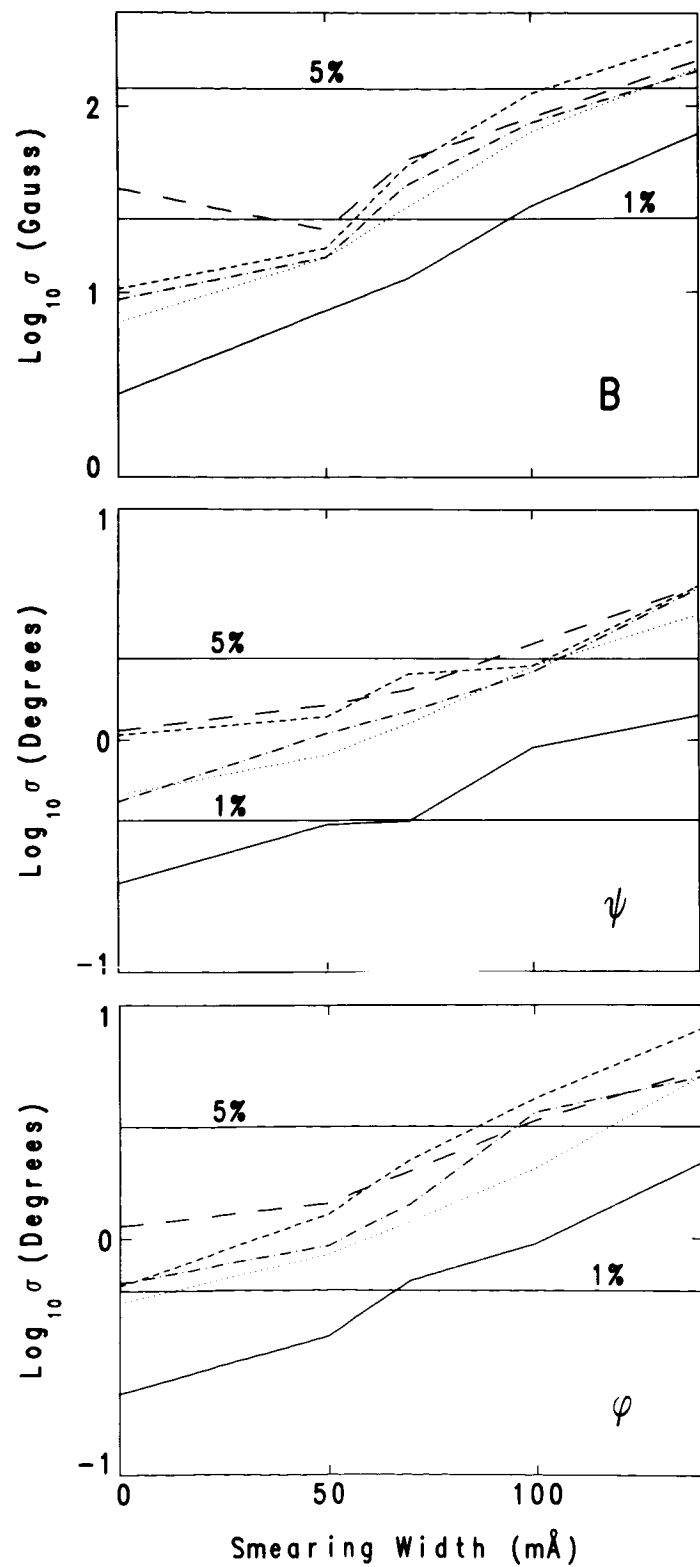


FIGURE 3. The rms error in the field expected for the standard case (see caption, Figure 1) is shown as a function of the Gaussian width of the filter function. The solid, dotted, dot-dashed, short dashed, and long dashed curves represent profile sampling intervals of 5, 25, 35, 50, and 70 mÅ, respectively. The horizontal lines again denote the 1% and 5% error levels for the field strength and angles.

of instrumental smearing for the sampling intervals listed in the caption. The 1% and 5% rms error levels are indicated by the horizontal lines. In some sense, the coarser sampling emulates the introduction of more noise in the profiles since fewer data points are used in the least-squares analysis. Figure 3 shows that, even with a rather coarse sampling, the rms error in the magnetic field strength and angles remains below 5% for 70 mÅ smearing width (FWHM = 116 mÅ). From this figure we see that accurate retrieval of the magnetic field may be obtained with this relative noise level with very broadband filters only if the line profile is sampled on a very fine grid. We have also investigated the systematic bias introduced into the inferred thermodynamic parameters by the smearing and coarse sampling. A statistically significant bias is introduced into η_0 by large smearing widths, but it probably results from the code limiting η_0 to positive values when the error in η_0 becomes significant relative to the adopted value of 10.

Vector magnetic fields in sunspots represent only a small proportion of potentially interesting solar magnetic field measurements. An example of an observational problem where weak fields need to be measured is the divergence and dilution of the quiet Sun magnetic field with height in the atmosphere. The polarization signal relative to the line intensity decreases with decreasing field strength, thereby decreasing the relative accuracy of the measurement. To illustrate this in a quantitative way we present the curves in Figures 4-7. We used the same atmospheric parameters as the standard case in the computations leading to Figure 2, however we varied the magnetic field strength from 250 to 3000 Gauss. The standard noise level of $\sigma = 0.003$ was adopted. The curves represent a selection of filters with the smearing widths given in the caption, and critically sampled (two sampling points per smearing width). The levels of 1% and 5% rms error are also indicated. Figure 4 shows that fields smaller than 1000 Gauss may not be reliably measured with filtergraphs without considerable reduction of the noise level.

To illustrate this problem more clearly, we present the results of a similar numerical experiment in Figures 5-7. For these computations we have chosen line parameters more closely approximating quiet Sun values ($\eta_0 = 10$, $\Delta\lambda_D = 40$ mÅ, $a = 0.1$, $\psi = 45^\circ$, and $\phi = 60^\circ$). Each of the panels in Figures 5-7 presents the rms noise level required of the measurement to achieve 5% rms error in the measurement (7% for Fig. 7) as a function of the sampling interval in mÅ. The separate curves of Figures 5-7 give the noise requirement for several values of the magnetic field. Figures 5-7 emphasize the necessity for considerable reduction of noise, presumably via averaging many images, in order to accurately recover the fields weaker than 1000 G.

The filter smearing widths for the computations leading to Figures 5-7 approximate the instrumental widths of several filters now being considered or used for this type of application. The narrowest filter (Fig. 5) approximates the Lockheed SOUP/SOT filter (Title and Torgerson 1983), the intermediate filter (Fig. 6) emulates the NASA/MSFC filter (Hagyard et al. 1982), and the filter with the widest passband (Fig. 7) has the approximate resolution of the Universal Birefringent Filter at the National Solar Observatory (NSO), (November

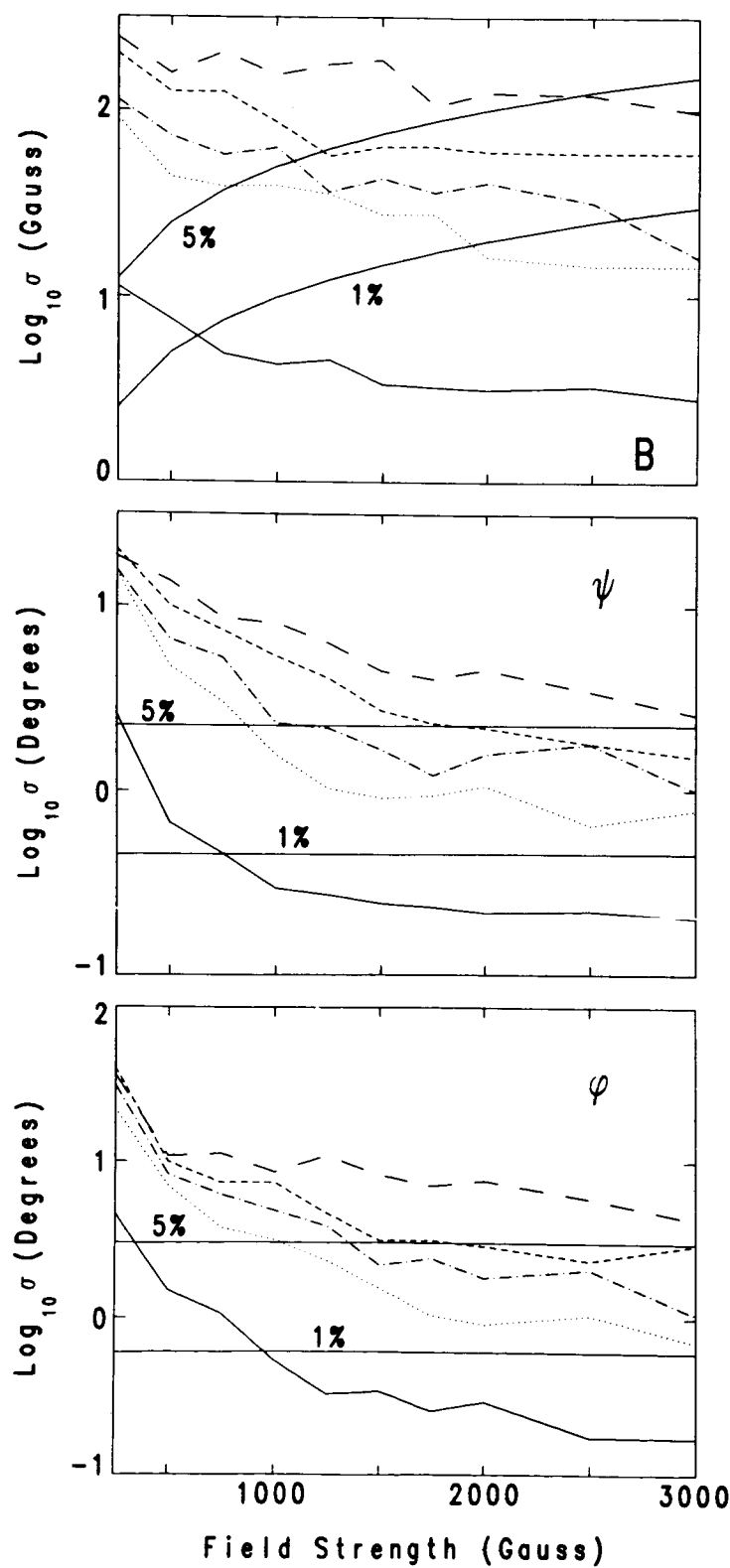


FIGURE 4. The rms error in the field parameters as a function of magnetic field strength are given for a noise level $\sigma = 0.003$. All parameters except the field strength are the same as those adopted for the standard case (see caption, Figure 1). The solid, dotted, dot-dashed, short dashed, and long dashed curves represent filters with smearing widths 0, 50, 70, 100, and 140 mA, respectively, and with sampling of the profiles at intervals equal to half the smearing width. Curves representing 1% and 5% rms error levels are also shown as solid lines.

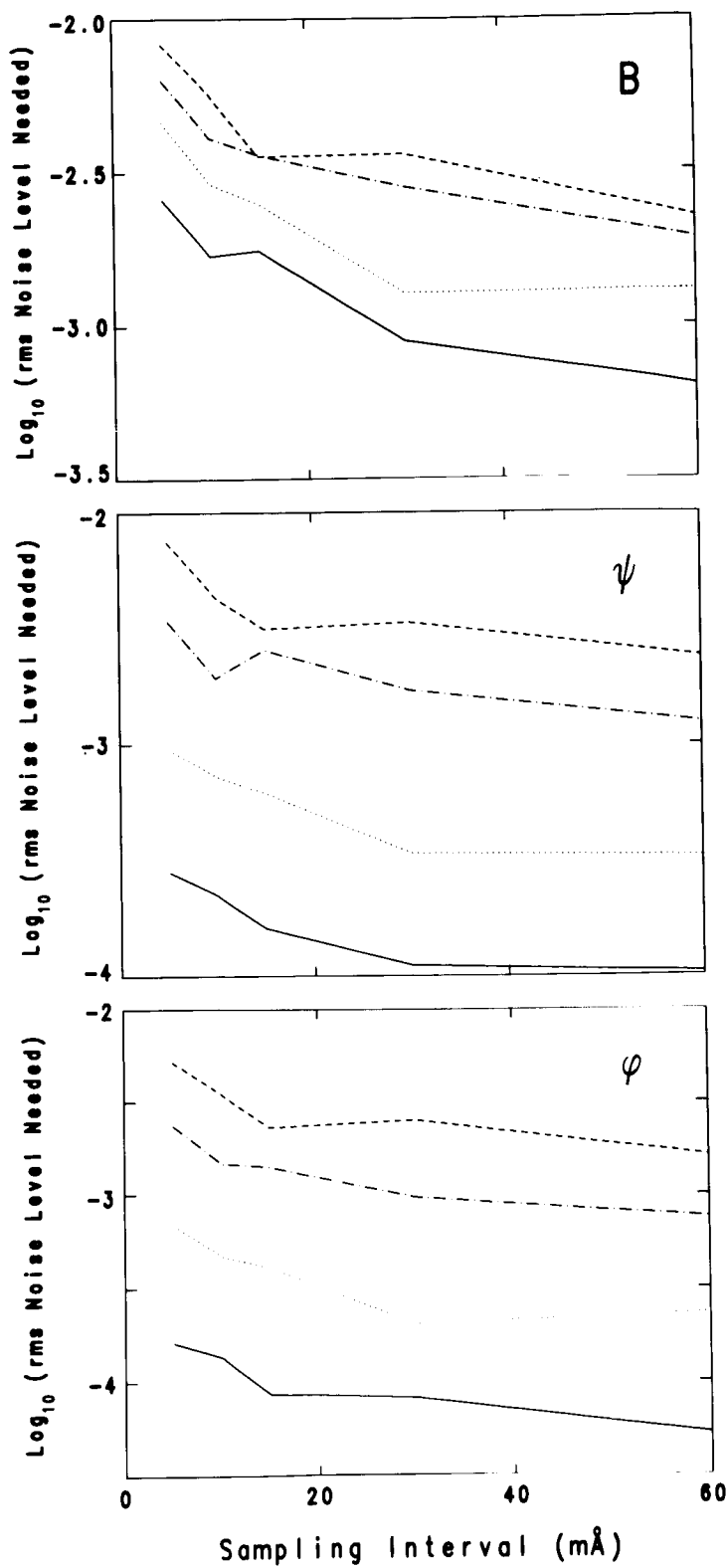


FIGURE 5. The rms noise level required of a measurement in order to obtain 5% accuracy in the magnetic field parameters is shown as a function of the sampling interval of the profile. The smearing width of the filter is assumed to be 30 $\text{m}\text{\AA}$ (50 $\text{m}\text{\AA}$ FWHM). The solid, dotted, dot-dashed, and dashed curves represent, respectively, field strengths of 100, 200, 400, and 600 Gauss. Field direction is fixed at $\phi=45^\circ$, $\psi=60^\circ$, and the assumed line parameters are $\eta=10$, $\Delta\lambda_D=40\text{m}\text{\AA}$, $a=0.1$.

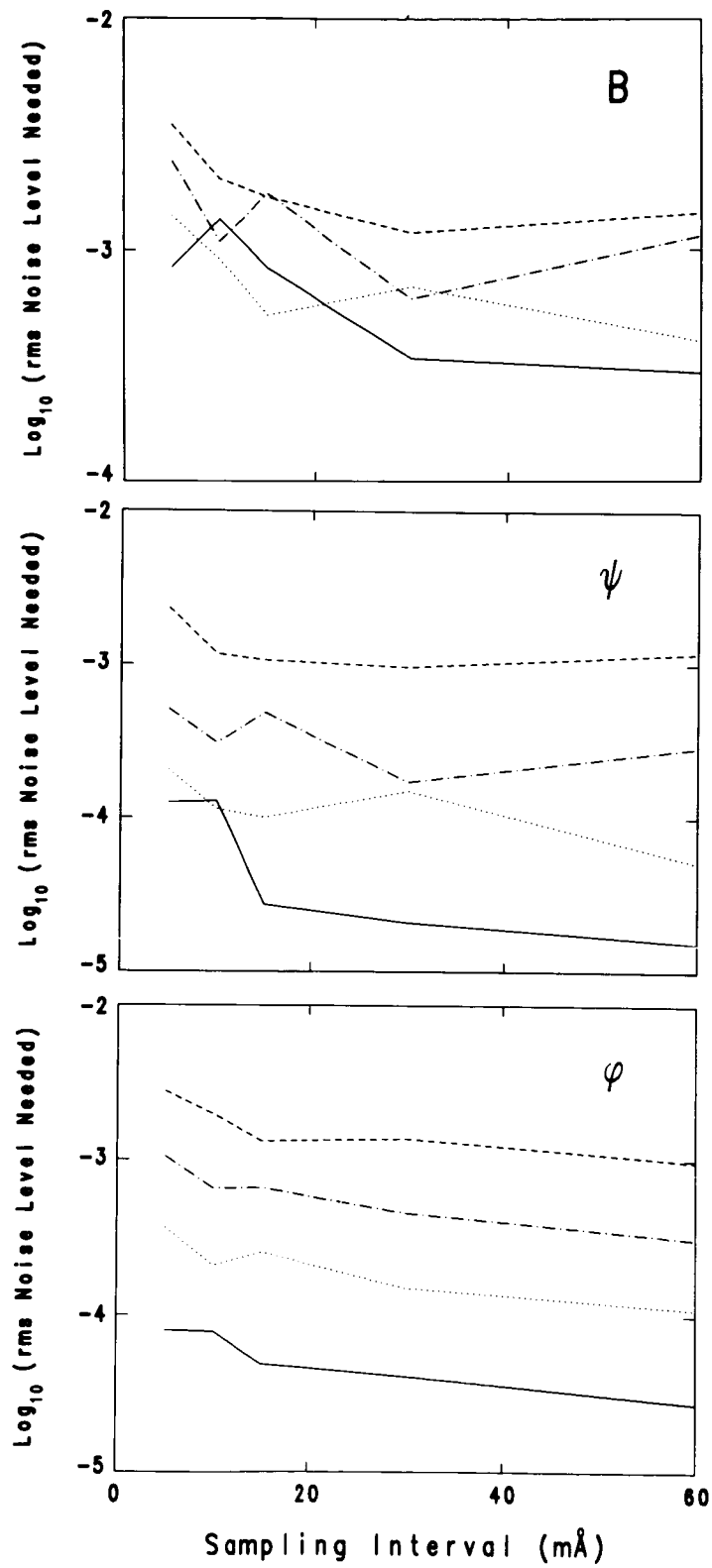


FIGURE 6. Same as figure 5, except the smearing width is assumed to be 75 mÅ (125 mÅ FWHM).

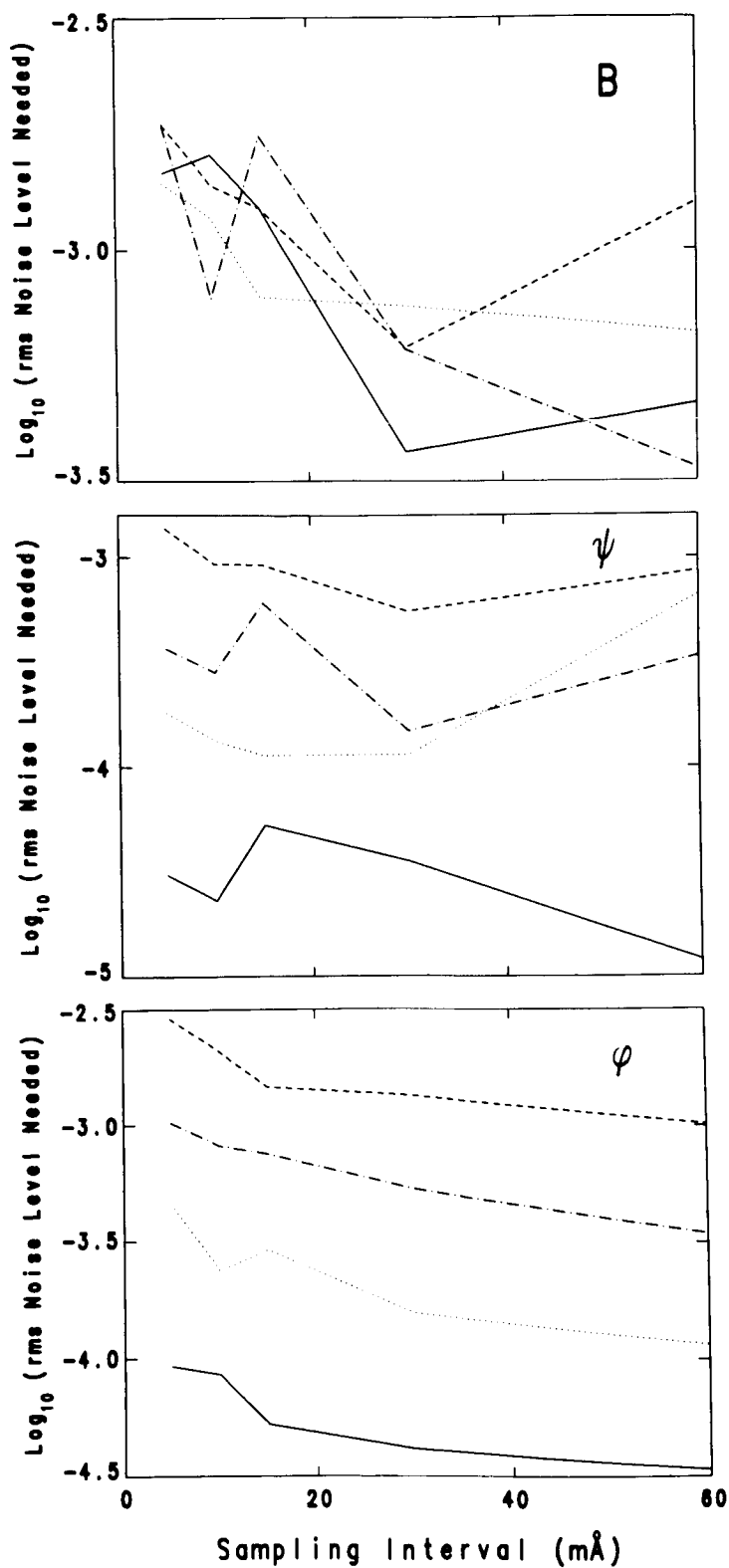


FIGURE 7. Same as Figure 5, except the smearing with is assumed to be 120 mÅ (200 mÅ FWHM), and the rms noise level shown is that required to achieve 7% accuracy.

and Stauffer 1984). Figures 5-7 should prove useful in future instrument and experiment design, as well as in assessing the accuracies available in current measurements.

The level of accuracy with which one may recover the field parameters is rather insensitive to reasonable changes in all the other parameters except the field inclination ψ . Figure 8 shows the behavior of the rms fluctuation in the field strength and orientation as a function of the field inclination. This computation was made with the standard parameters and curves are shown for the same selection of smearing width and critical sampling used above in Figure 4. Obviously, the field angles are best determined when ψ is near 90° since the linear polarization signal is strongest there. We also note that the analysis imposes the restriction $0.1^\circ < \psi < 179.9^\circ$, so that low field inclinations are biased toward higher values, and high field inclinations are biased toward lower values. The 180° ambiguity in the azimuthal angle ϕ becomes crucial as the magnetic field is nearly aligned with the line-of-sight near $\psi = 0^\circ$ or 180° in Figure 8. At field inclinations so close to the line-of-sight, the azimuthal angle plays a negligible role.

IV. SYSTEMATIC EFFECTS

A. Errors in the Adopted Filter Profile

Our inversion technique incorporates the instrument function directly, so any difference between the adopted and true instrument function will result in some error in the magnetic field and thermodynamic parameters. We first investigate the systematic errors introduced by errors in the wings of the instrument function. The Stokes vectors were convolved with Voigt functions instead of Gaussians, then they were analyzed on the basis of the Gaussian instrument function alone. We have used our standard strong field case (see Section III) with the same selection of Gaussian smearing widths and critical sampling used for the computations leading to Figure 4. The promising results are given in Figure 9. We see very little influence on the derived parameters for Voigt parameters of the instrument function $a < 0.1$, and even for $a = 0.3$ the errors introduced are generally modest when compared with the errors introduced by the choice of sampling and smearing. There seems to be little effect on the derived value of ϕ for this choice of parameters even for large values of the Voigt "error." Figure 9 shows that one may expect an increase in the rms error of ϕ of about a factor of 2 with $a = 0.3$, independent of the smearing width and sampling interval. The magnitude of the magnetic field seems most strongly influenced by this systematic error at small smearing widths. The zero smearing width case, of course, is not influenced by a . The apparent fluctuations in this case give a measure of the repeatability of different realizations. The introduction of the Voigt "error" in the assumed instrument function naturally introduces a bias in the derived line damping parameter and μB_1 , but the other line parameters are little changed. This numerical experiment tells us that small errors in the wings of the instrument function may be easily tolerated.

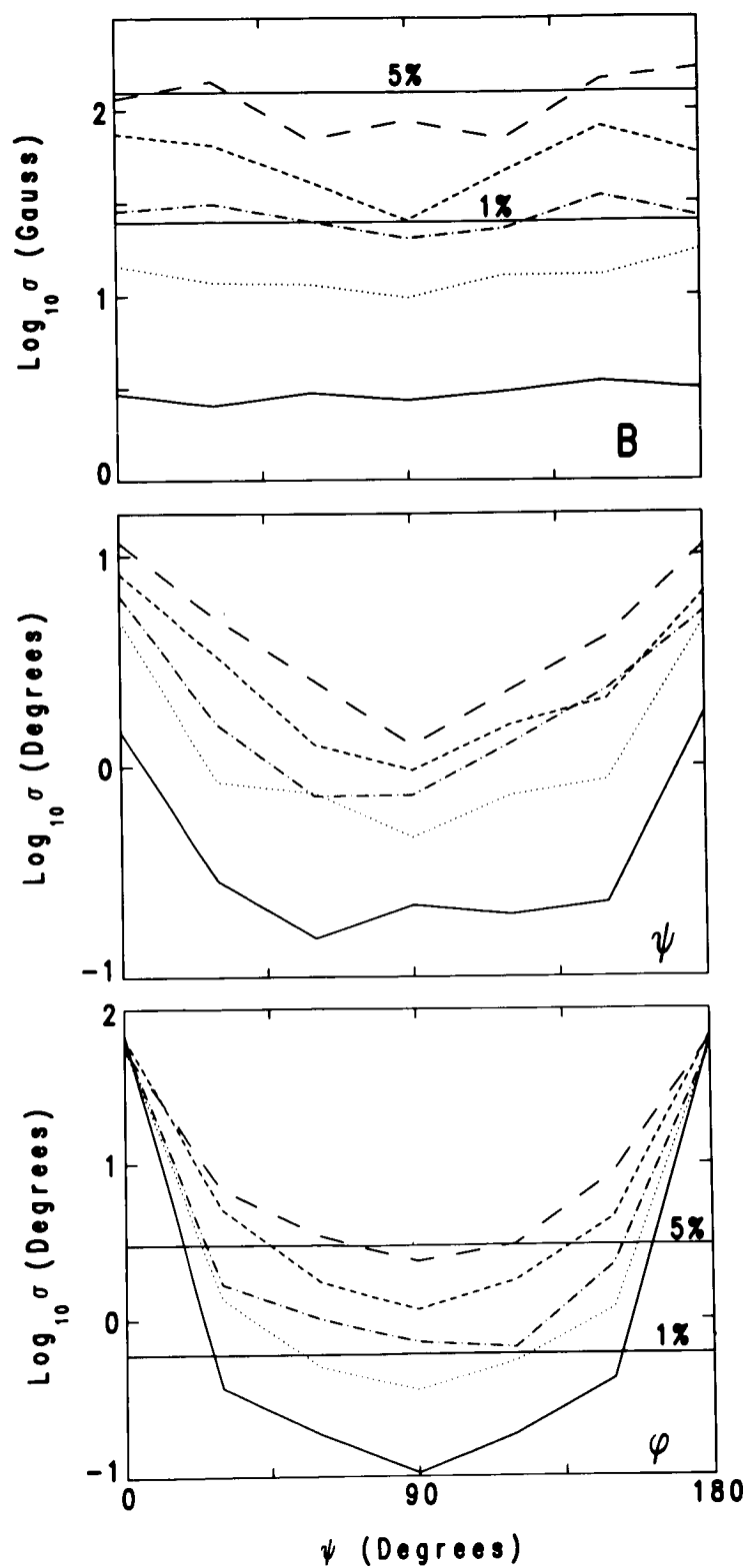


FIGURE 8. The rms noise level expected in the magnetic field strength and orientation is shown as a function of the field inclination ψ with respect to the line-of-sight. The standard case (see caption, Figure 1) with $\sigma = 0.003$ was used, and the curves show different smearing widths and sampling intervals, which are identical to those used in Figure 4. The horizontal lines in the upper and lower panels represent rms error levels of 1% and 5%.

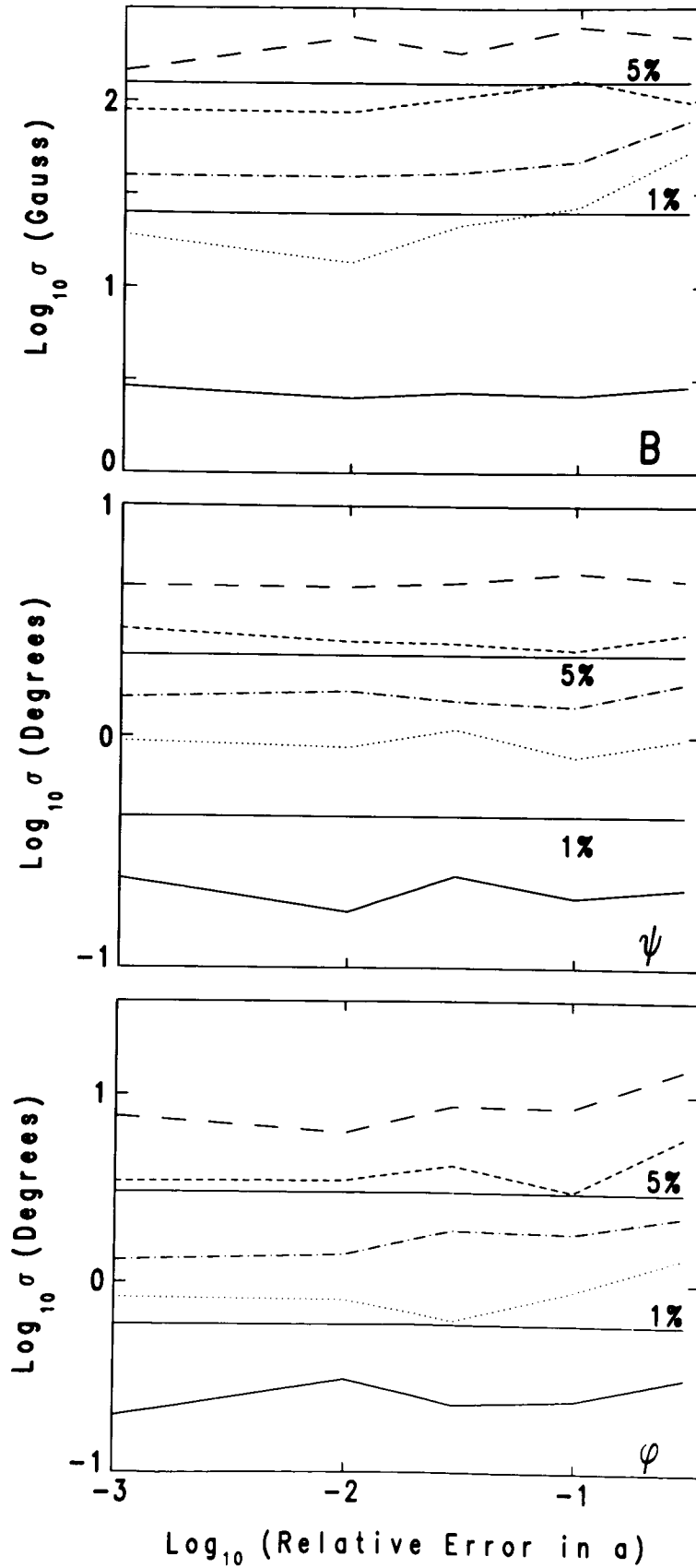


FIGURE 9. The rms noise levels expected in the magnetic field parameters as a function of the strength of the wings of the true filter wavelength resolution function are shown. The profiles were smeared with a Voigt function having a damping parameter a (shown on the abscissa) and then analyzed on the basis of an adopted Gaussian filter function. The curves show different smearing widths and sampling intervals, which are labeled in an identical fashion to Figure 4. The solid horizontal lines represent rms error levels of 1% and 5%.

We have also checked the effects of errors in the Gaussian width of the adopted instrument function for the strong field case. We introduced errors in the adopted instrument function of $\pm 20\%$, and find results similar to the wing "errors" case described above. The field inclination ϕ is insensitive to the width of the adopted instrument function, whereas errors in the field azimuth ϕ increase by about a factor of 2 with a 20% error in the smearing width. The error introduced into the field strength by underestimates of the smearing width is greater than that introduced by overestimates. Narrow filters are again most seriously affected by this systematic error. For example, the error in the field strength introduced by a 20% underestimate of the width of a 50 mÅ filter is a factor of 4 larger than the error expected with the correct filter width. A 20% overestimate of this filter width increases the error by only a factor of 1.6.

We have not made an exhaustive parameter study of this type of systematic error, so the weak field case may be more sensitive to uncertainties in the instrument function. Additionally, asymmetric errors in the adopted instrument function may introduce much larger errors in the inferred field strength and direction than the results above would indicate. Our results show that some error in the knowledge of the wings of the instrument function may be tolerated, but the width of the primary lobe of the instrument function should not be underestimated.

B. The Influence of Solar Oscillations

It is not possible to sample a line profile instantaneously with a filter-graph, so, relative to the actual instantaneous line profiles, solar oscillations will cause a distortion of profiles measured over a finite time interval. In the photosphere the dominant 5-minute oscillations are evanescent, so the entire photospheric column responds to these modes by moving in an essentially vertical manner as a unit (Lites and Chipman 1979). The first-order effect on the line profiles is thus a wavelength shift of the entire line profile. The 5-minute oscillations are reduced greatly in amplitude in sunspot umbrae relative to the quiet Sun, and they are diminished in amplitude to some degree in plage areas and in sunspot penumbrae. We thus expect the oscillations to be most troublesome to the measurement of magnetic fields in the quiet Sun, less of a problem in strong field regions, and very little of a problem in umbral photospheres. We have chosen a test case that emulates a sunspot penumbra or a plage to demonstrate the errors that can be expected in measurement of magnetic fields during a finite fraction of an oscillation cycle. The parameters adopted here are the same as those of the standard case except that the field strength is 1500 Gauss. As a worst case, we assume that the amplitude of the oscillation is ± 0.5 km/s (photospheric amplitudes are usually smaller than this) and we determine the errors introduced to the fit by sampling the profile over fractions of the oscillation cycle. The phase of the sinusoidal oscillation is assumed to be zero at the start of the observation in these simulations. We compare the fits to the Stokes profiles distorted by oscillations to the fits to the undistorted but otherwise identical profiles in order to isolate the effects of the oscillations. Figure 10 presents the errors introduced into the magnetic field parameters by the oscillatory distortion of the profiles for the usual selection of smearing widths

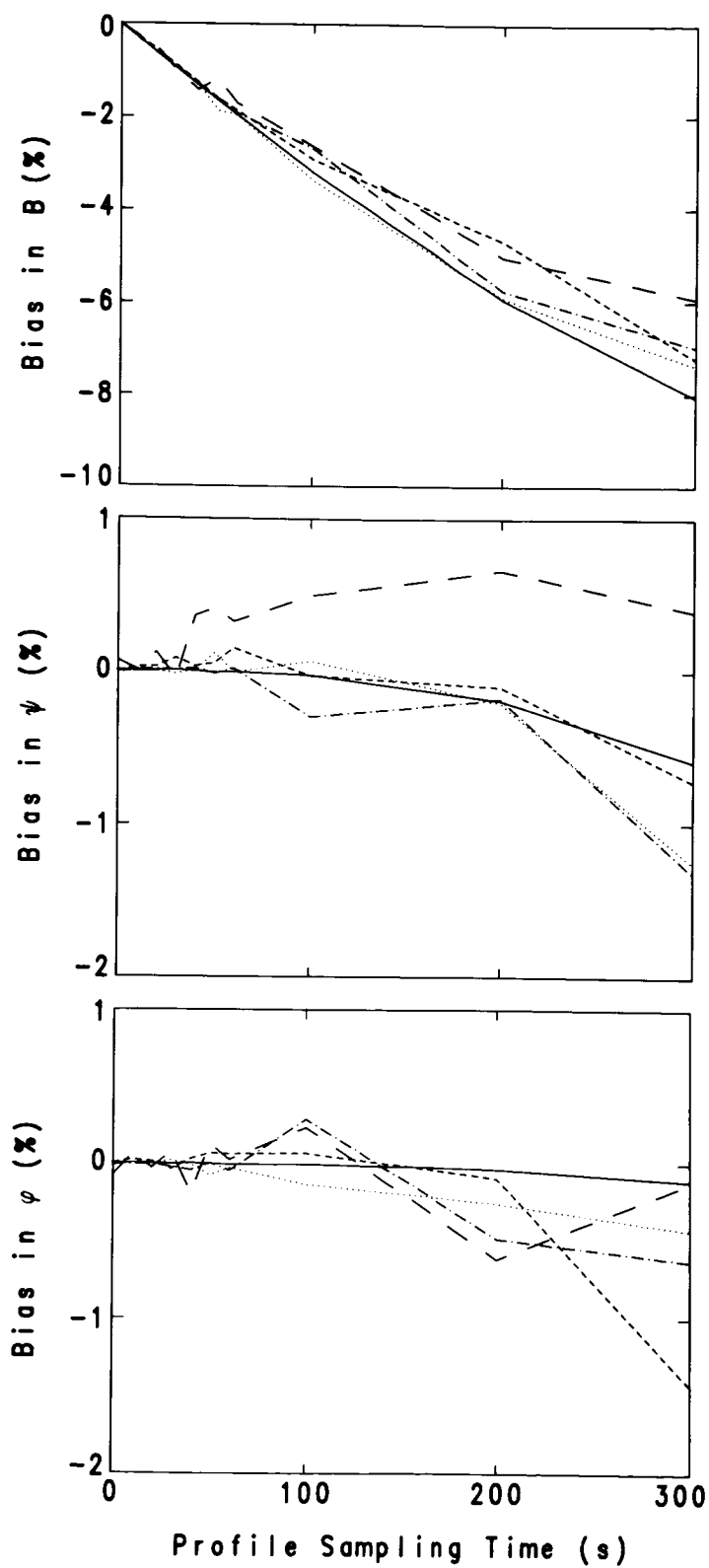


FIGURE 10. The bias in the magnetic field parameters introduced by a sinusoidal shift in the line profile of 300 second duration is shown as a function of sampling time for the profile. The noise level adopted is $\sigma = 0.003$, although the bias shown compares the fits to the undistorted profiles (with the same noise as introduced into the undistorted profiles) with the fits to the distorted profiles. The standard strong field case was modified to a 1500 Gauss field to simulate a sunspot penumbra or plage area. Curves for different sampling intervals, smearing widths are as labeled in Figure 4.

with critical sampling. This figure shows that the oscillations introduce a systematic bias into the derived field strength, as one might expect, since it causes a net change in apparent separation of the σ components of the line profile. The worst case changes the derived field strength by less than 8%. The gross effect is on the magnetic field strength, since the oscillation alters the field angles by only about 1% in the worst case. Figure 10 shows that the bias in the magnetic field strength is essentially independent of the adopted field strength and sampling interval.

We have not tested the influence that oscillations may have on the weak field case, but one should expect the relative errors to be larger there. On the other hand, the results from the previous section indicate that one may be forced to average many measurements together in order to reduce the error to reasonable levels. We expect that such an averaging will also reduce the influence of the oscillatory distortion through time averaging.

V. FALSE SOLUTIONS AND FAILURE TO CONVERGE

In any non-linear least-squares fitting procedure there is always the danger that the scheme will not converge, or converge to a secondary minimum far from the correct solution. These problems may be minimized by any or all of three routes: good initial guesses for the field and thermodynamic parameters, reduction of the noise level of the observations, and use of narrow filters with fine wavelength sampling of the line profiles. For only two of the over 10^4 sets of Stokes profiles considered in the above analyses did this method yield a magnetic field that was obviously well outside the statistical variance of the other realizations of the noise. In cases where rms level of noise is greater than 0.01, one encounters rather large errors in the derived parameters more frequently, and such errors must be considered a limitation of the analysis. As mentioned above, the method converges (albeit slowly) to the correct solution even when the initial guess is far removed from it. Several ways may be envisioned to locate and compensate for poor solutions. Among those that first come to mind are: to check for failure to converge in a reasonable number of iterations, to compare the solution for consistency with adjacent spatial points, and to compare several vector magnetic field maps taken in rapid succession. We do not feel that the infrequent occurrence of a false solution derived from polarization images of adequate wavelength sampling and noise level will be a fundamental limitation to this method.

One may pose another question related to convergence: are there certain combinations of angles, field strengths, filter widths, and sampling intervals that will, in the presence of noise, lead the least-squares fitting routine to a false solution far from the actual solution? In order to check this, we have simulated a series of measurements of a sunspot from near disk center to close to the limb, and we find no consistent problems of this type. We adopt the magnetic field strength and angle at the approximate height where Fe I $\lambda 6173$ is formed from the model of Landman and Finn (1979). Table 1 lists our values of magnetic field strength, angle of inclination, and η_0 that we have interpolated as a function of the fraction of the penumbral radius R_p . The umbral radius is about 0.35

R_p in this model. The spot is assumed to be radially symmetric. We also assume that the Doppler width and damping are constant at 35 mÅ and 0.1, respectively. The gradient of the source function is assumed constant over the spot, so we adjust the parameter μB_1 according to the disk position of the simulated sunspot. Finally, we have adopted a noise level in the continuum of 10^{-3} for quiet Sun outside the sunspot and assume that the observed intensities in the penumbra and umbra are 0.7 and 0.15 the quiet Sun intensity, so that the rms noise in the penumbral and umbral areas are 1.2×10^{-3} and 2.58×10^{-3} , respectively.

We calculate the magnetic field angles in their axially symmetric spot model for the 21 radial positions listed in Table 1, for each 20° sector of the axial angle of the spot, and for disk angles (from disk center) of 1° , 15° , 30° , 45° , and 60° . We assume that we can make a very good guess for the field for the center of the sunspot, and derive the least-squares fit to the parameters at that point. From the center we work our way outward along radii using the parameters derived for the previous point as an initial guess for the least squares procedure. We have carried out this procedure for filters with 30, 75, and 120 mÅ resolution and critical sampling as described above. These choices of filters simulate existing Lockheed, NASA/MSFC, and NSO filters, respectively.

The results shown as images of the error in the derived magnetic field parameters in Figures 11 - 13 show that there are no regions apparent where our program systematically finds a false solution. The most obvious feature of these images is the consistent region where the 180° uncertainty of the azimuthal angle of the field is apparent. In an observation like this, it would be very easy to remove the ambiguity of the field azimuth since the program behaves in such a consistent fashion. We also note a few points near the edge of some of the images in which the error in the inferred field parameters are much larger than usual. This problem may be anticipated from the combination of the low field strength in the outer regions of the penumbra and the adopted noise level (see Section III). These isolated instances could be interpolated over after the fact. We expect that real measurements will often cover regions where the magnetic field changes strength and orientation rapidly with spatial position, so a better means of handling those cases would be to make three or more measurements of the same area in rapid succession. The infrequency of these discrepant points would surely allow one to select the correct field from three measurements.

VI. DISCUSSION

This study shows that one may recover the magnetic field strength and direction from filtergraph images of the solar surface obtained in both linear and circular polarization within magnetically-sensitive line profiles. However, there are some inherent limitations to this method. The accuracy of the inversion of the data decreases with the decreasing spectral resolution of the filter. Good results may be expected with high resolution filters now available. Recovery of weak vector fields will require very high quality data. For these weak fields, very narrow filters will be required with excellent signal-to-noise and fine sampling through the profile. For example, to recover the direction of a 100 Gauss field within 5% using a 50 mÅ filter, a signal-to-noise ratio of 10^4 will

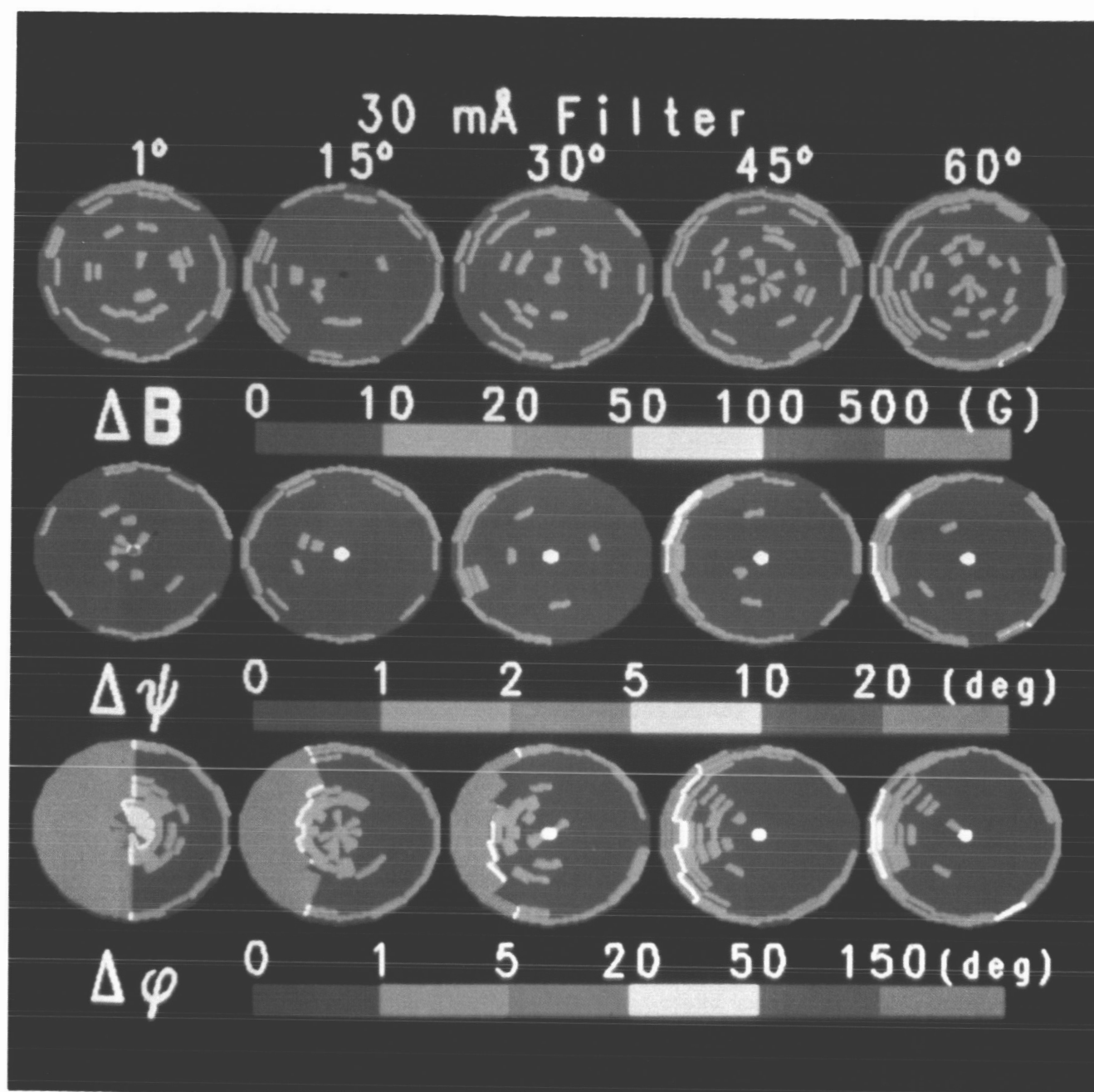


FIGURE 11. Color-coded errors in the field strength and angles are shown for a filter of width 30 mÅ (50 mÅ FWHM) and 15 mÅ sampling interval. Each column represents the model sunspot displaced from disk-center by the number of degrees shown. The areas of each of the 361 radial elements in this simulated observation are adjusted (by varying the radial extent) so that all have equal areas.

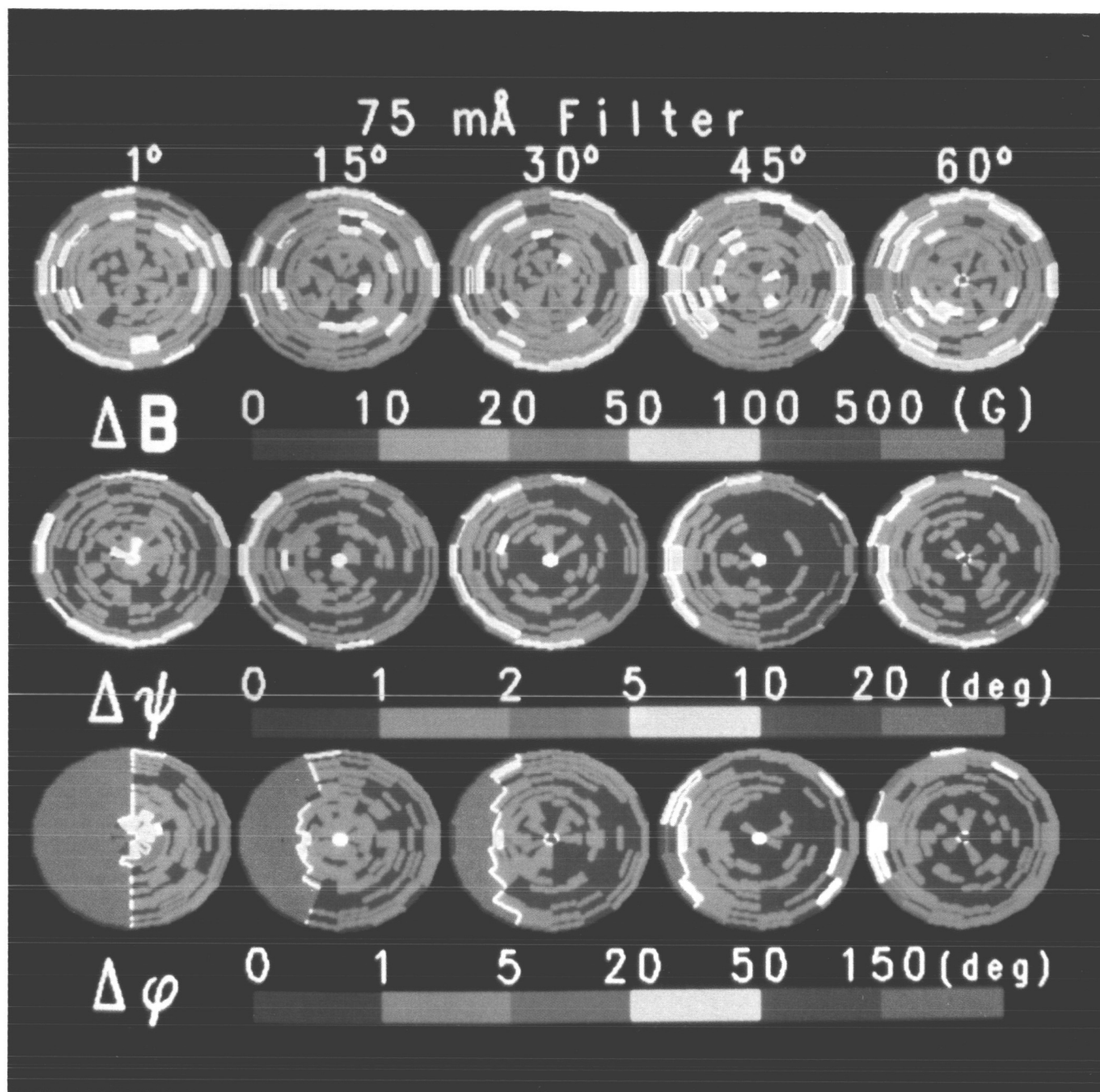


FIGURE 12. Same as Figure 11, except a filter of width 75 mÅ (FWHM 125 mÅ) and 37.5 mÅ sampling interval was used.

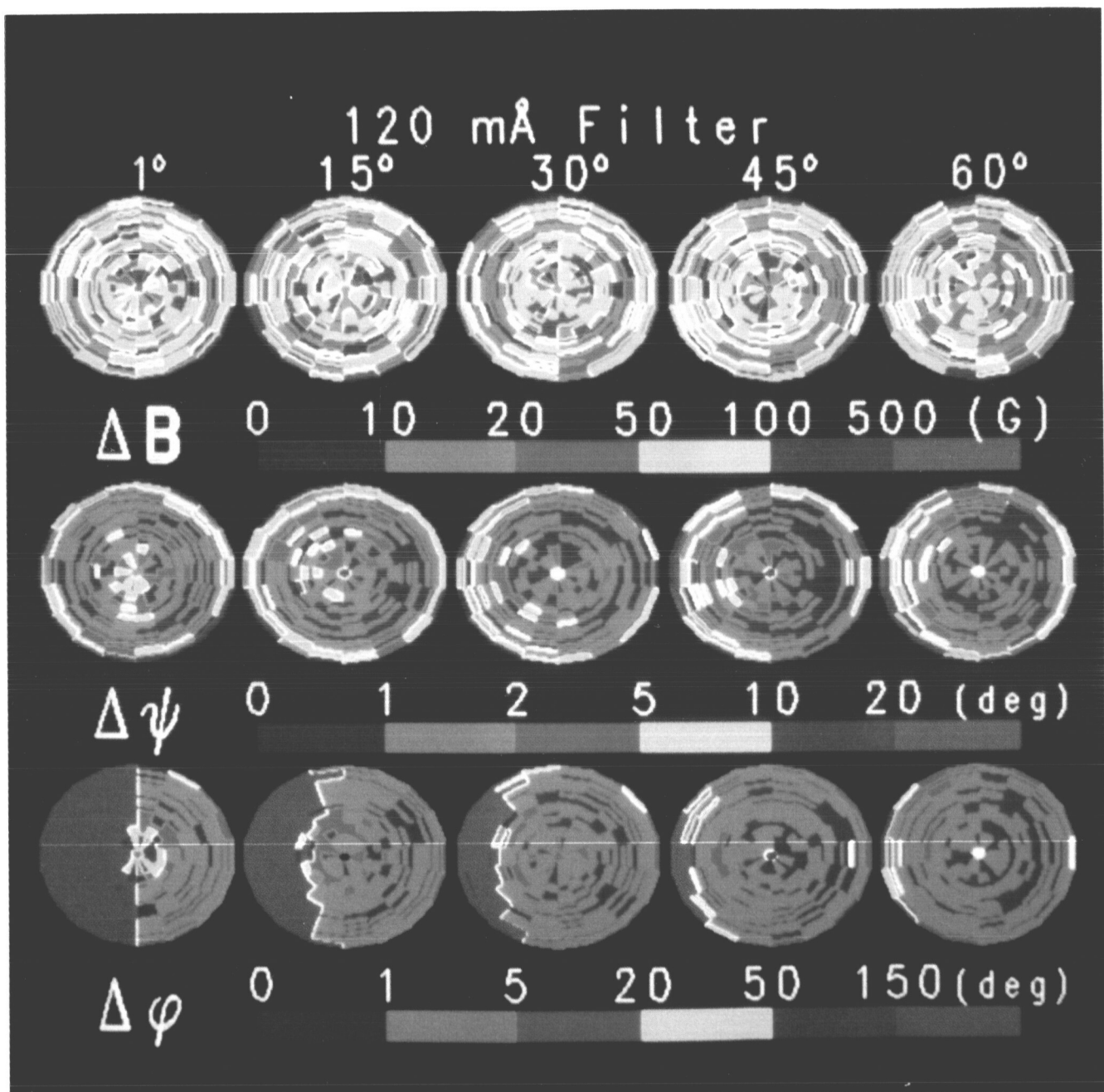


FIGURE 13. Same as Figure 11, except a filter of width 120 mÅ (FWHM 200 mÅ) and 60 mÅ sampling interval was used.

be needed. Weak vector field maps will be very difficult to come by. There are several other limitations we have not considered that may prove to be significant sources of error to this kind of analysis. One of these is residual depolarization and crosstalk between the Stokes parameters that cannot be removed from the data during the calibration process. In order to be consistent with the random noise level of the dataset σ , one should hold this form of error to $\sigma n^{-1/2}$, where n is the number of sampling points through the profile. This will keep the random offset of the polarization signals to about the same amount as the systematic error. For the weak field case, this limitation could be very difficult to achieve.

The limitations of the Unno/Rachkovsky theory are also of some concern. Non-uniform magnetic fields within the line formation region, gradients in the thermodynamic parameters that depart from the assumptions, and gradients in the line-of-sight motions may all lead to some error in the mean field within the line formation region. These limitations are discussed separately in Skumanich, Rees, and Lites (1984). In the chromosphere this method will generally not be applicable because the line source functions are often non-monotonic functions of line optical depth.

When the field of view of an element within the image does not resolve rapid horizontal changes in the magnetic field, one may expect our routine to fail in unexpected ways. This is a fundamental limitation of the analysis that may, in general, only be resolved by higher spatial resolution. Numerical investigation of the limitation of spatial resolution may be a subject of further study. Imaging polarimetry has some potential for partial restoration for the instrumental and atmospheric spatial degradation. As such, the possibility of partial correction should be considered an advantage of using filtergraphs to do polarimetry.

Finally, we have demonstrated that the thermodynamic parameters are not as well determined as the magnetic field strength and direction. The accuracy of the inversion procedure and the computational burden may be improved by using simultaneous measurements of a non-magnetic line to establish some of the thermodynamic parameters $\Delta\lambda_D$, μB_1 , and perhaps even η_0 .

We wish to thank Dr. R. R. Fisher for several illuminating discussions and suggestions.

REFERENCES

- Auer, L. H., Heasley, J. N., and House, L. L.: 1977, Solar Phys. **55**, 47.
- Brault, J. W. and White, O. R.: 1971, Astron. Astrophys. **13**, 169.
- Hagyard, M. J., Cumings, N. P., West, E. A., and Smith, J. E.: 1982, Solar Phys. **80**, 33.
- Landman, D. A. and Finn, G. D.: 1979, Solar Phys. **63**, 221.
- Lites, B. W. and Chipman, E. G.: 1979, Ap. J. **231**, 570.
- Lites, B. W. and Skumanich, A.: 1984, "Description and Operation of the Routines CHISQMO, INVRTS, VOIGEN, VOIGTV, SMEAR, RUFF, and POSCAR in the Stokes Analysis Program" SPO Internal Memorandum T2-84.
- November, L. J. and Stauffer, F.: 1984, (Submitted to J. Opt. Soc. Am.).
- Rachkovsky, D. N.: 1962, Izv. Krymsk. Astrofiz. Obs. **27**, 148.
- Rachkovsky, D. N.: 1967, Izv. Krymsk. Astrofiz. Obs. **37**, 56.
- Rayrole, J.: 1982, in "Annual Report, Joint Organization for Solar Observations" (A. v. Alvensleben, Ed.), Kiepenheuer-Institut fur Sonnenphysik: Freiburg, West Germany.
- Skumanich, A. and Lites, B. W.: 1984 (in these proceedings).
- Skumanich, A., Rees, D., and Lites, B. W.: 1984 (in these proceedings).
- Title, A. M. and Torgerson, D. D.: 1983, B.A.A.S. **15** (No. 2), 709.
- Unno, W.: 1956, Publ. Astron. Soc. Japan **8**, 108.

TABLE 1.

Sunspot magnetic field strengths and angles are listed as interpolated from the model of Landman and Finn (1979). R_p is the penumbral radius, γ is the inclination of the field, and η_0 is the ratio of the line center to continuum opacity. Values listed are for the approximate level when the line is formed. The umbra-penumbra boundary is at $r/R_p \approx 0.35$.

r/R_p	$ B $ (Gauss)	γ (degrees)	η_0
0	3460	0	16.0
0.05	3440	1.4	15.6
0.10	3400	5.3	14.4
0.15	3320	10.5	12.9
0.20	3220	15.5	10.7
0.25	3070	20.1	8.5
0.30	2910	24.4	7.2
0.35	2730	27.9	7.0
0.40	2500	31.0	7.0
0.45	2210	33.0	7.0
0.50	1980	36.0	7.0
0.55	1740	38.1	7.0
0.60	1520	40.3	7.0
0.65	1350	42.5	7.0
0.70	1250	44.8	7.0
0.75	1050	46.7	7.0
0.80	900	48.1	7.0
0.85	750	49.9	7.0
0.90	600	51.6	7.0
0.95	400	53.5	7.0
1.00	300	55.2	7.0

SOLAR MONOCHROMATIC IMAGES IN MAGNETO-SENSITIVE SPECTRAL LINES AND MAPS OF VECTOR MAGNETIC FIELDS

YE SHIHUI, JIN JIEHAI, and JIANG MINHAN

Purple Mountain Observatory, Academia Sinica

Abstract. This paper describes a new method which allows us by use of the monochromatic images in some magneto-sensitive spectral line, e.g. FeI λ 6302.499, to derive both the magnetic field strength as well as the angle between magnetic field lines and line of sight for various places in solar active regions. In this way two-dimensional maps of vector magnetic fields may be constructed. We have tentatively applied this method to our own observational material and some reasonable results have been obtained. In addition, we have worked out a project for constructing the three-dimensional maps of vector magnetic fields.

I. Introduction

One of the main tasks of contemporary solar observations is to get the maps of the magnetic fields of the sun, particularly in active regions. The measurement of longitudinal magnetic fields cannot completely satisfy the requirements of solar research. It is desirable to have vector magnetic field maps characterized by three quantities: the magnetic field strength (H), the angle between field lines and the line of sight (γ), and the azimuthal angle (χ). However, the measurements of the magnetic vector have been confronted with many difficulties. For instance, the vector magnetograph of the Crimean Astrophysical Observatory (Stepanov and Severny, 1962; Severny, 1964) has continued to use the principle of the photo-electric magnetograph of H. W. Babcock (1953), i.e., the measurements of magnetic fields with the modulation at line wings. This method has the following shortcomings: (1) The exit slits are narrow and so the photo-electric signals are weak, while the amplification with large ratios may give rise to distortion. (2) If two exit slits are used, it is difficult to put them at exactly symmetric positions. (3) The two wings of the magneto-sensitive line may be asymmetric in themselves. For example, the red wing of FeI λ 6302.499 is interrupted by two blending lines (Wittmann, 1972). (4) In the case of strong magnetic fields the saturation effect may be remarkable. (5) Babcock's photo-electric magnetograph can measure only the longitudinal component of the field strength (H_{\parallel}), while the vector magnetograph of Stepanov and Severny—only the transverse component (H_{\perp}). In the former case $\gamma = 0$, while in the latter case $\gamma = \pi/2$. In both cases one cannot know the value of the angle γ for the total field.

Owing to the technical difficulties in making magnetographs, as well as the drawbacks stated above, many astronomers have tried to supplement magnetic maps with monochromatic images. There has been some controversy about the relationship between $H\alpha$ images and magnetic maps, yet it is quite definite that there is a strong correlation between the monochromatic images of CaIIK and CN λ 3883 and magnetic fields (Ye, 1978). But it has to be pointed out that these monochromatic images cannot provide quantitative information of magnetic fields. This is because these lines do not exhibit the normal Zeeman effect and so the existing theory of spectral line formation can hardly be applied to them.

Several years ago R. E. Loughhead and R. J. Bray (1976, 1978) proposed a method with which one may derive the strength and direction of magnetic fields according to the contrast of magnetic and non-magnetic regions on the monochromatic images of FeI λ 5250.216. But they use Unno's algebraic solution of the equations of transfer of Stokes parameters (Unno, 1956) and we have shown (Ye, et al., 1978) that this solution may give rise to large errors. Moreover, they have performed only theoretical calculations without application to observational data. It should also be mentioned that the line λ 5250 is very sensitive to temperature and in recent years it has been discarded in the observations of solar magnetic fields.

The purpose of this work is to establish a convenient method for the derivation of vector magnetic field maps from the monochromatic images in a certain magneto-sensitive line. Its merits and defects will be described in the third section.

II. The relation Between the Equivalent Widths of a Magneto-Sensitive Line and Magnetic Fields

In our work (Ye, et al., 1978) we found a numerical solution to the equations of transfer for Stokes parameters in the presence of magnetic fields (i.e., Unno's equations). It may be used to derive the equivalent widths of the magneto-sensitive line used and then the intensities of the monochromatic images in this line. As we shall show a little later, it is better to use the monochromatic images of both ordinary ("o") and extraordinary ("e") polarized light. For the computations of the "o" and "e" line profiles we use the following formulae (Chandrasekhar, 1950):

$$\begin{aligned} I_{X,\delta} &= \frac{1}{2} (I - Q \sin 2\phi \cos \delta + U \cos 2\phi \cos \delta + V \sin \delta) , \\ I_{Y,\delta} &= \frac{1}{2} (I + Q \sin 2\phi \cos \delta - U \cos 2\phi \cos \delta - V \sin \delta) . \end{aligned} \quad (1)$$

where X and Y represent the directions of the "o" and "e" radiations respectively and ϕ is the angle formed by the fast axis of the 1/4-wavelength plate and the X-axis. For the moment we take the third Stokes parameter $U=0$. For the 1/4-wavelength plate the phase difference $\delta = \pi/2$. Then the above formulae become

$$\begin{aligned} I_{X,\lambda/4} &= \frac{1}{2} (I + V) \\ I_{Y,\lambda/4} &= \frac{1}{2} (I - V) . \end{aligned} \quad (2)$$

Taking the intensity of the continuum to be unity, the intensities in line profiles may be changed to residual intensities. Then Eqs. (2) may be re-written as follows:

$$\begin{aligned} r_o &= r_1 + r_V \\ r_e &= r_1 - r_V . \end{aligned} \quad (3)$$

By use of our numerical solution of Unno's equations and with various values of γ and H we have calculated several "o" and "e" profiles of λ 6302 and a few examples can be seen in Figs. 1 and 2. Only "o" profiles

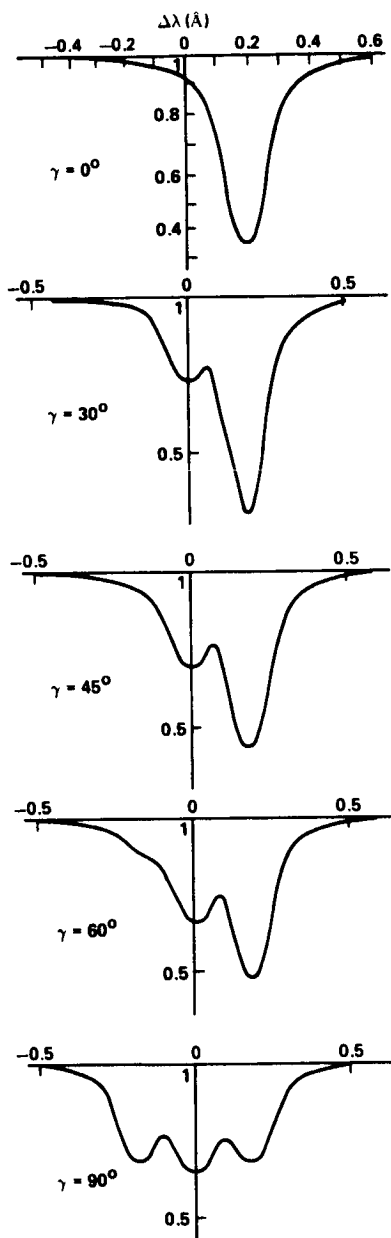


Fig. 1. The "o" profiles of FeI λ 6302.499 with $H=4000\text{G}$ and $\gamma = 0^\circ, 30^\circ, 45^\circ, 60^\circ, 90^\circ$.

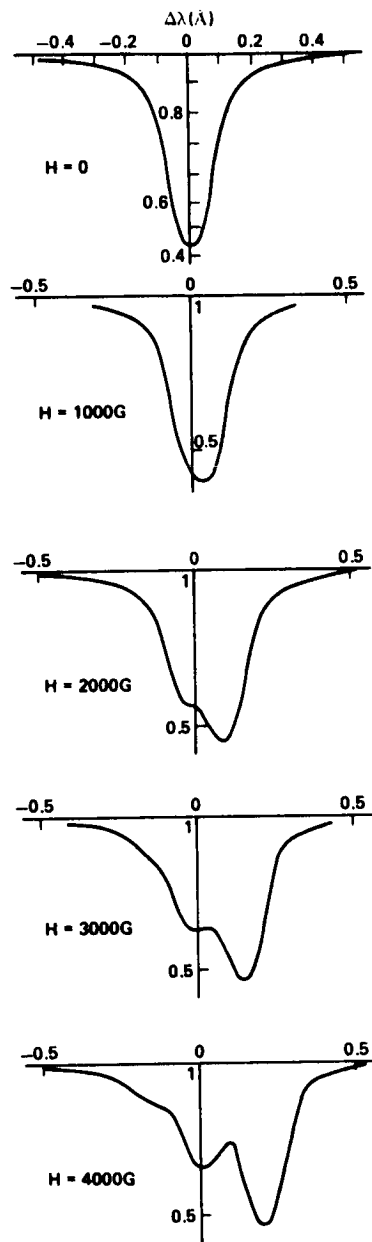


Fig. 2. The "o" profiles of FeI λ 6302.499 with $\gamma = 70^\circ$ and $H = 0, 1000, 2000, 3000, 4000\text{G}$.

are shown, because the "o" and "e" profiles are inverses of each other with respect to line center. From these figures it may be readily seen that line profiles and hence equivalent widths W_λ change remarkably with both H and γ . Therefore, it is possible in principle to derive H and γ from W_λ , and thus from the intensities on monochromatic images.

In the scanning of monochromatic images both the width ($\Delta\lambda$) and the position of the slit of the monochromator with respect to line profile (a,b) may be chosen freely (see Fig. 3). For a definite slit width and position the intensity of the monochromatic image within a spectral line is defined as $I(\Delta\lambda) = \int_a^b i_\lambda d\lambda$ and the intensity of the monochromatic image in continuum is given by $I_c(\Delta\lambda) = i_c\Delta\lambda$. Then we have the following relation:

$$\frac{I(\Delta\lambda)}{I_c(\Delta\lambda)} = 1 - \frac{W_\lambda(\Delta\lambda)}{\Delta\lambda} \quad (4)$$

where $W_\lambda(\Delta\lambda) = \int_a^b \frac{i_c - i_\lambda}{i_c} d\lambda$ is the equivalent width

of the "o" or "e" magneto-sensitive line obtained under a specific condition of the entrance slit. This quantity may be calculated with the theoretical profiles given by our numerical solution of Unno's equations as well as Eqs. (3). As for the intensity on the monochromatic image in a magneto-sensitive line in units of the intensity of the same point on the monochromatic image in continuum, i.e., $I(\Delta\lambda)/I_c(\Delta\lambda)$, it may be determined with observational material. Since the "o" or "e" profiles are closely connected with H and γ , both $W_\lambda(\Delta\lambda)$ and $I(\Delta\lambda)$ should also depend on these two parameters. This is the basis on which we may derive H and γ from the intensities of any one and the same point on the monochromatic images.

In order to express the relationship between $W_\lambda(\Delta\lambda)$ (in the following abbreviated as W) and H as well as γ , we have adopted three different fashions. The first is the graphic representation. An example can be seen in Fig. 4 which exhibits the dependence of W on H and γ for $a = -0.2\text{\AA}$, $b = 0$.

The second is the approximate expressions. For the above slit condition and with $\gamma = \pi/4$ we have established with a Chinese-made TQ-6 electronic computer the following expression of H in terms of W :

$$\begin{aligned} 10^{-3}H = & 1.32 \times 10 - 2.60 \times 10^2 W - 1.51 \times 10^3 W^2 + 1.02 \times 10^4 W^3 \\ & + 3.52 \times 10^5 W^4 + 4.36 \times 10^8 W^5 + 2.35 \times 10^7 W^8 \\ & - 3.43 \times 10^8 W^7 - 1.39 \times 10^{10} W^8 \end{aligned} \quad (5)$$

Thirdly, we may use the method of interpolation. We have calculated with the same computer two tables and each of them pertains to one of the following slit configurations: (1) $a = -0.2\text{\AA}$, $b = 0$; (2) $a = -0.1\text{\AA}$, $b = 0.1\text{\AA}$. In each table a series of values of H and γ are arranged according to the increasing order of the values of W . The H values are confined in the range from 0 to 4000G with a constant interval of 50G. The γ 's are in the range $0 - 90^\circ$ and the interval of neighboring values is 5° . In each table there are 1539 pairs of H and γ values. Due to the shortage of space, these tables cannot be contained in this paper. As an illustration five groups of values are given in the following small table. All these values belong to the first slit configuration.

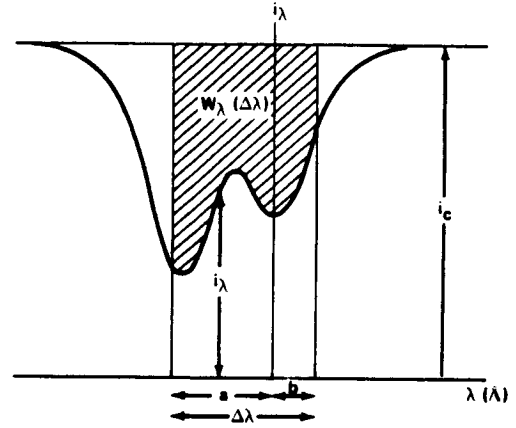


Fig. 3. Relation between the intensity of monochromatic image and the equivalent width of a spectral line.

Table 1. The relation between W and γ as well as H.

$10^2 W (\text{\AA})$	$\gamma (^{\circ})$	H (G)
3.705	40	2250
3.709	15	1000
3.711	10	950
3.715	30	1400
3.719	25	1200

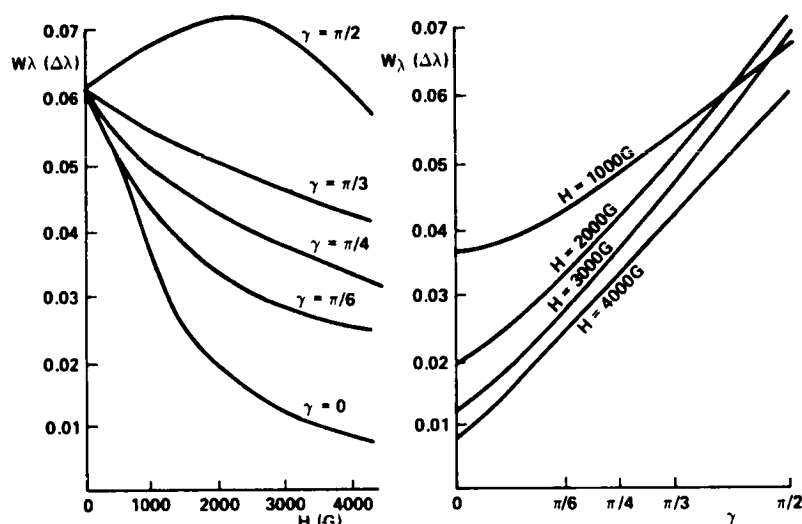


Fig. 4. The dependence of W on H and γ .

Each of the above three methods has its own advantages and disadvantages. After a careful consideration we prefer the third description of the variation of W with H and γ .

III. Derivation of Vector Magnetic Maps from Monochromatic Images in a Magneto-Sensitive Line

It has been stated in the foregoing section that the intensity of each point on a monochromatic image in a magneto-sensitive line is related to H and γ . In order to determine these two physical quantities we are in need of the two values $I(\Delta\lambda_1)$ and $I(\Delta\lambda_2)$, i.e. the intensities for one and the same point on monochromatic images scanned with the two slit configurations $\Delta\lambda_1$ and $\Delta\lambda_2$. All three of the above-mentioned descriptions of variation of W with H and γ are the direct or indirect reflection of the following functional relations:

$$\begin{aligned} W_{\lambda}(\Delta\lambda_1) &= f_1(H, \gamma) \\ W_{\lambda}(\Delta\lambda_2) &= f_2(H, \gamma) \end{aligned} \quad (6)$$

Since these functions are known already, they can be substituted into Eq. (4) and we have

$$\frac{I(\Delta\lambda_1)}{I_c(\Delta\lambda_1)} = 1 - \frac{f_1(H, \gamma)}{\Delta\lambda_1} \quad (7)$$

$$\frac{I(\Delta\lambda_2)}{I_c(\Delta\lambda_2)} = 1 - \frac{f_2(H, \gamma)}{\Delta\lambda_2}$$

By measuring the intensities at the same point on the two monochromatic images we get the ratios on the left sides of these equations. Then by solving them we may determine both H and γ for the given point. If we repeat this procedure for many points, we can draw a vector magnetic field map. In reality, we resort to the third method stated above and hence no computation is needed. It is only necessary to look through the tables.

It has been pointed out in the first section that both the classical method of H. W. Babcock and the new method of Loughhead and Bray have a series of shortcomings. Our way of derivating the magnetic vector with monochromatic images in magneto-sensitive lines can avoid the saturation effect and some other defects. But, it must be noted that we treat only the line profiles of polarized radiation "o" or "e". This is because the influence of the changes of H and γ on these profiles are much more remarkable than on the profiles of unpolarized or natural light, and hence the former are more profitable in magnetic field measurements. Loughhead and Bray considered only unpolarized radiation and in this respect our method is also an improvement of theirs. However, it should be emphasized that our method has its own drawbacks. First, it requires that the variations of the intensities on monochromatic images are caused solely by magnetic fields, while the effects due to the changes of temperature, density and other parameters are negligible. This is a supposition which of course does not coincide with reality. But we have shown (Ye et al., 1978) that the temperature sensitivity of λ 6302 is small. Therefore, at least for this line the above-mentioned question is not serious. Secondly, in Unno's equations there is the factor $\cos\theta$ (θ is the angle between the normal on the solar surface and the line of sight) and θ is different for different points on the solar surface. For the sake of simplicity one may take a mean $\cos\theta$ for the scanned region and this may give rise to errors. Yet for the central region of the solar surface $\cos\theta$ does not vary by much and so these errors may be small. Thirdly, our method can be applied only to regions of strong magnetic fields. This is because weak fields may have only small influences on line profiles so their effects on monochromatic images are difficult to detect.

IV. A Tentative Observation of the Vector Magnetic Field of a Sunspot

We have applied the method established in this paper to our own observational material. The results show that both H and γ may be determined with monochromatic images.

Specifically, we have scanned a F-type spot group in the central region of the solar disc (with heliocentric longitude 9° and latitude $+19^\circ$) on November 3, 1979 with the monochromator of the Purple Mountain Observatory and have secured the monochromatic images in γ 6302 with the above-mentioned two configurations of the entrance slit. For the largest spot in this group there were two umbrae in one penumbra. In the larger umbra, i.e. the one on the western side, we have measured the intensities at the center on the two images and got $I(\Delta\lambda_1)/I_c(\Delta\lambda_1) = 0.80$, $I(\Delta\lambda_2)/I_c(\Delta\lambda_2) = 0.79$. Substituting these values into Eqs. (5) and (6) we have the equivalent widths $W_\lambda(\Delta\lambda_1) = 0.026 \text{ \AA}$,

$W_{\lambda}(\Delta\lambda_2) = 0.042 \text{ \AA}$. Then from our tables we have found $H = 3650\text{G}$, $\gamma = 25^\circ$. Besides this, in the outward direction from the umbra center and at three points with angular distances 3.5, 7.0 and 10.5 arcsecs the following results have been obtained: $H = 3200\text{G}$, $\gamma = 45^\circ$; $H = 2800\text{G}$, $\gamma = 50^\circ$ and $H = 3700\text{G}$, $\gamma = 60^\circ$ respectively. These data roughly agree with the commonly adopted magnetic field strengths of sunspots as well as the law of the distribution of their field lines. It is still more important to note that for every point we have gotten almost the same values of H and γ with the two W 's. Although our instrument is not highly accurate and there may be large observational errors, the tentative observation described above shows that our method is applicable.

V. Determination of the Three-Dimensional Magnetic Vector

In order to construct three-dimensional magnetic field maps we need not only H and γ but also the azimuthal angle of the magnetic field line χ . In our work (Ye et al., 1978) we followed Unno's simplifying assumption that the third Stokes parameter U is equal to zero. According to the following formula:

$$\tan 2\chi = \frac{U}{Q}, \quad (8)$$

we in this case always have $\chi = 0$. Then for any point on the solar surface one may get only two-dimensional maps. In order to construct three-dimensional maps we cannot take $U = 0$ and must solve the system of equations of transfer of the four Stokes parameters I , Q , U and V (i.e., Unno-Beckers' equations, in Beckers 1969). We have done the numerical solution in Ye and Jin (1983). The intensity of radiation I thus obtained must be a function of H , γ and χ . Moreover, the equivalent widths should also be functions of these three quantities, i.e.

$$W_{\lambda}(\Delta\lambda) = f(H, \gamma, \chi) \quad (9)$$

Hence if we scan the same area on the solar surface with three configurations of the entrance slit $\Delta\lambda_1$, $\Delta\lambda_2$ and $\Delta\lambda_3$ and for the same point measure the intensities on the three corresponding monochromatic images, we may by a formula similar to Eq. (4) get the three equivalent widths $W_{\lambda}(\Delta\lambda_1)$, $W_{\lambda}(\Delta\lambda_2)$ and $W_{\lambda}(\Delta\lambda_3)$. With the known functional relations (9) we may simultaneously obtain the three quantities H , γ and χ from the three W_{λ} 's. If we repeat this procedure for many points, we may construct three-dimensional magnetic field maps.

References

- Babcock, A. B. 1953, *Astrophys. J.*, 118, 387.
- Beckers, J. M. 1969, *Solar Phys.*, 9, 372.
- Chandrasekhar, S. 1950, "Radiative Transfer".
- Loughhead, R. E., Bray, R. J. 1976, *Solar Phys.*, 50, 297.
- Loughhead, R. E., Bray, R. J. 1978, *Publ. Astr. Soc. Pac.*, 90, 230.
- Severny, A. B., Izv. Krymskoj 1964, *Astrofiz. Obs.*, 31, 126.
- Stepanov, V. E., Severny, A. B., Izv. Krymskoj 1962, *Astrofiz. Obs.*, 28, 166.
- Unno, W. 1956, *Publ. Astr. Soc. Japan*, 8, 108.
- Wittmann, A. 1972, *Solar Phys.*, 23, 294.
- Ye Shihui 1978, "The Magnetic Fields of Celestial Bodies", *Science Press*, 6.2 and 5.1.
- Ye Shihui, Wang Zhenyi, Jin Jiehai 1978, *Acta Astronomica Sinica*, 19, 152.
- Ye Shihui, Jin Jiehai 1983, *Acta Astronomica Sinica*, 24, 119.

MAGNETIC FIELD-VECTOR MEASUREMENTS IN QUIESCENT PROMINENCES VIA THE HANLE EFFECT: ANALYSIS OF PROMINENCES OBSERVED AT PIC-DU-MIDI AND AT SACRAMENTO PEAK

V. Bommier,* J. L. Leroy,** and S. Sahal-Br      *

**Observatoire de Paris-Meudon, 92190 Meudon, France*

***Observatoire du Pic-du-Midi et de Toulouse, 31400 Toulouse, France*

The Hanle effect method for magnetic field vector diagnostics has now provided results on the magnetic field strength and direction in quiescent prominences, from linear polarization measurements in the He I D₃ line, performed at the Pic-du-Midi and at Sacramento Peak. However, there is an inescapable ambiguity in the field vector determination: each polarization measurement provides two field vector solutions symmetrical with respect to the line-of-sight. A statistical analysis capable of solving this ambiguity has been recently applied to the large sample of prominences observed at the Pic-du-Midi (Leroy et al., 1984); the same method of analysis applied to the prominences observed at Sacramento Peak (Athay et al., 1983) provides results in agreement on the most probable magnetic structure of prominences; these results are detailed in the following paragraphs.

The Sacramento Peak spectropolarimeter has provided narrow-band polarization data, in which the two components of He I D₃ are fully resolved, which form a pair of lines adequate for the achievement of the complete field vector determination (Bommier et al., 1981). In a sample of 14 prominences observed and interpreted (Athay et al., 1983), the two symmetrical field vectors have been found to approximate 90° starting from the local solar radius, i.e. neighbouring the horizontal plane.

Two-line observations in broad-band polarization have been recently performed with the Pic-du-Midi coronagraph polarimeter, in He I D₃ and Hydrogen H  , which have given the same result on the field direction (Leroy, Bommier, Sahal-Br      , to be published). Considering that result, one-line observations previously made of the unresolved He I D₃ line have been interpreted in terms of the field strength and azimuth angle only in 120 prominences of the Polar Crown (Leroy et al., 1983) and 256 prominences of medium and low latitude (Leroy et al., 1984).

In most cases, the two symmetrical field-vectors do not cross the prominence long axis in the same sense; further investigation is then required before we are able to provide observational constraints on the type of the magnetic structure of prominences, in particular with respect to the polarity of the photospheric magnetic field on each side of the neutral line underlying the prominence.

Owing to the symmetry of the two field vectors with respect to the line-of-sight, one has the following "mirror effect" within the two angles α_V and α_F between the field vectors and the prominence long axis, and the angle β between the prominence long axis and the line-of-sight (Fig. 1):

$$\alpha_V + \alpha_F = 2\beta \quad (1)$$

The precise definition of the sign and value of α and β can be found in Leroy et al. (1984, see Fig. 4 of that paper); our convention is that positive and negative α angles correspond to prominence field orientations consistent and inconsistent respectively with the potential and non-potential field orientation with respect to the polarity of the neighbouring photospheric field.

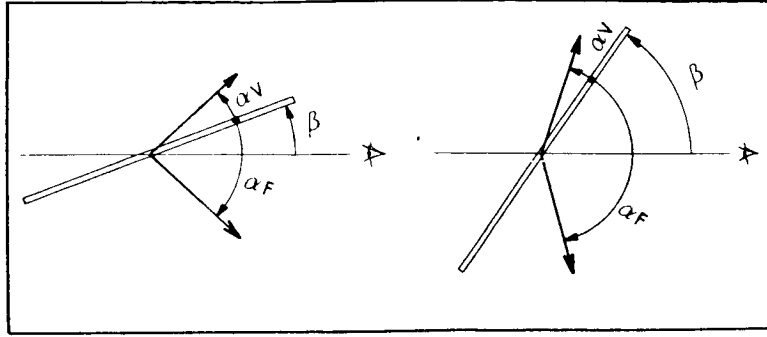


Fig. 1. The magnetic field determination through a polarimetric analysis provides two field vector solutions symmetrical with respect to the line-of-sight.

If there is a preferential value α_c in the sample of prominences, one has:

$$\alpha_V = \alpha_c \quad (2)$$

then according to (1)

$$\alpha_F = -\alpha_c + 2\beta \quad (3)$$

For a couple of two observed field vectors, there is no a-priori choice of α_F and α_V between the two α angles.

Figure 2 is a plot of relations (2) and (3) for $\alpha_c = 30^\circ$. The points on Fig. 2 correspond to the observed α and β values for 12 of the 14 prominences observed at Sacramento Peak. Two prominences have been discarded, as the direction of the prominence long axis cannot be established. The values of the α angles have been derived by averaging the values obtained for each measurement by Athay et al., (1983, see table 2 of that paper). In Table 1 and Fig. 2, the prominence names refer to that paper; in each prominence, the averages have been done for each day and each position angle.

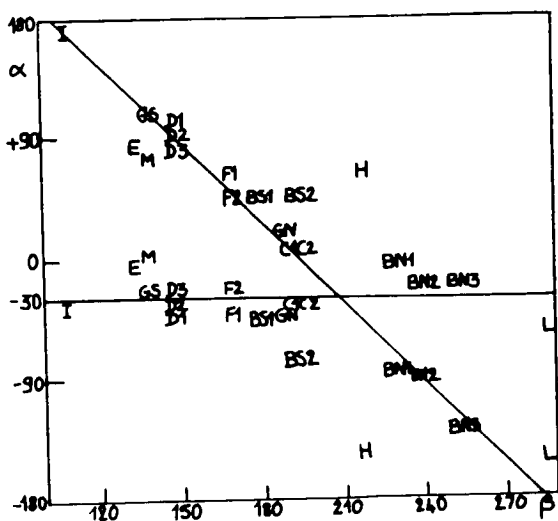


Fig. 2. Plot of relations (2) and (3) for $\alpha_c = -30^\circ$, and of the observed averaged values for 12 prominences observed at Sacramento Peak (Athay et al., 1983). Compare with Fig. 6 of Leroy et al. (1984).

A similar plot can be found on Fig. 6 of Leroy et al. (1984), for 120 of 256 prominences observed at Pic-du-Midi, for which the geometrical parameters and the direction of the underlying neutral line, can be established with accuracy. This sample of 120 prominences corresponds to 900 measurements.

Figure 2 represents the least-square fitting of (2) and (3) in the observed data; the squared residues $\Sigma (0 - C)^2$ in the least-square fitting are given on Figure 3 as a function of α_c , for the Pic-du-Midi data (full line, Fig. 7 of Leroy et al., 1984), and for the Sacramento Peak data (dotted line). In the fitting calculation, the data must be weighed according to their accuracy. Two kinds of data uncertainties have to be considered:

- 1) Geometrical errors, which are roughly constant over one given prominence.
- 2) Photometric errors which can occur on each measurement.

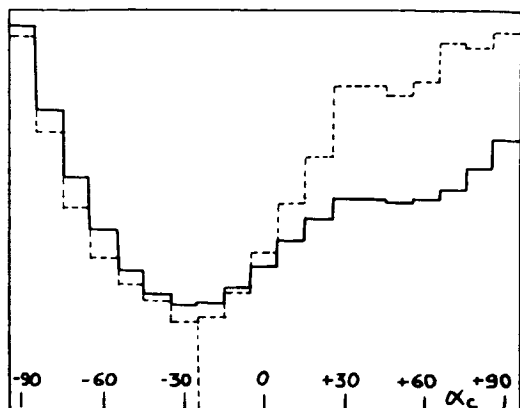


Fig. 3. The squared residues in the least-squares fitting of relations (2) and (3) in the Pic-du-Midi data (full line) and in the Sacramento Peak data (dotted line).

In the Pic-du-Midi data, a strict selection of 120 prominences out of 256 ones has been done in order to eliminate the geometrical errors; then, for all individual measurements the weights have been determined according to photometric accuracy only. In the Sacramento Peak data, the geometrical errors cannot be discarded in the same way, owing to the smaller size of the prominence sample; it is then not realistic to weigh each prominence, in which geometrical errors can occur, according to the number of averaged measurements $w = \nu$; on the other hand, it would not be statistically fair to give the same weight $w = 1$ to all the prominences because in some of them one measurement only has been done, for which photometric uncertainty is likely. As a compromise, we have used the values $w \sim \sqrt{\nu}$ listed in Table 1 as weights in the least-square fitting of the Sacramento Peak data. However, the least-square fittings with $w = 1$ and $w = \nu$ have also been done, and have given the same minimum $\alpha_m \sim -30^\circ$ for the squared residues $(O - C)^2$ as the fitting given on Fig. 3.

The most probable value $\alpha_m \sim -30^\circ$, which is obtained as the result of the fitting, corresponds to the non-potential orientation of the prominence magnetic field with respect to the polarity of the neighbouring photospheric field. The good agreement which can be seen on Fig. 3 between the fits in the two sets of data shows that the same trend can be found in the two sets of prominences observed by the two instruments. However, the large size of the Pic-du-Midi sample of prominences has enabled more detailed analysis which have given evidence to two types of prominences (Leroy et al., 1984):

1) Prominences with maximum height larger than $\sim 30\,000$ km. Their magnetic structure is consistent with the Kupperus-Raadu type of models (i.e. non-potential sense of the prominence field). The mean α value is $\sim 25^\circ$ and the mean field strength is ~ 5 to 10 Gauss. These prominences often show filamentary or curtain-like structures. Polar Crown prominences fall in this category.

2) Prominences with maximum height lower than $\sim 30\,000$ km; their magnetic structure is consistent with the Kippenhahn-Schlüter type of model (i.e. potential sense of prominence field). The mean α value is $\sim 20^\circ$ and the mean strength is ~ 20 Gauss. These prominences are bright, often sharp-edged in He I D_3 and found essentially at low latitude.

These statistical results have been confirmed on favourable individual cases: for 15 prominences observed at Pic-du-Midi, the two-field vectors are pointing on the same side of the prominence, and the α angles are large enough with respect to the measurements and interpretation inaccuracies, so that the field polarity is derived without any ambiguity.

References

- Athay, R. G., Querfeld, C. W., Smartt, R. N., Landi degl'Innocenti, E., Bommier, V.: 1983, *Solar Phys.*, **89**, 3.
- Bommier, V., Leroy, J. L., Sahal-Bréchet, S.: 1981, *Astron. Astrophys.*, **100**, 231.
- Leroy, J. L., Bommier, V., Sahal-Bréchet, S.: 1983, *Solar Phys.*, **83**, 135.
- Leroy, J. L., Bommier, V., Sahal-Bréchet, S.: 1984, *Astron. Astrophys.*, **131**, 33.
- Leroy, J. L., Bommier, V., Sahal-Bréchet, S., to be published.

Table 1. Averaged values for 12 prominences observed at Sacramento Peak (Athay et al., 1983) of the prominence aspect angle AA and azimuth angle ϕ^* and ϕ_s^* of the two magnetic field solutions with respect to the prominence long axis. The angles AA, ϕ^* , ϕ_s^* are respectively related to the angles β , α_F , α_V defined in Leroy et al. (1984). ν is the averaged number of measurements.

Prominence	Position Angle	AA	β	ϕ^* , ϕ_s^*	α_F , α_V	ν	$w \sim \sqrt{\nu}$
BN1 79/04/24	212	-50	-130	-177, -83	-3, -97	4	2
BN2 79/04/25	212	-60	-120	-159, -81	-21, -99	1	1
BN3 79/04/26	215	-75	-105	-159, -52	-21, -128	2	1
BS1 79/04/26	208	0	-180	+135, -135	+45, -45	3	2
BS2 79/04/27	209	-15	-165	+132, -102	+48, -78	1	1
C1 79/04/25	27	15	-165	-36, +6	-36, +6	1	1
C2 79/04/27	28	15	-165	-35, +5	-35, +5	1	1
D1 80/08/15	121	30	150	+80, -140	+100, -40	20	4
D2 80/08/15	114	30	150	+87, -147	+93, -33	46	7
D3 80/08/16	115	30	150	+94, -154	+86, -26	19	4
E 80/09/17	262	45	135	+89, -179	+91, -1	1	1
F1 80/09/17	355	-10	-10	+70, -50	+70, -50	4	2
F2 80/09/19	355	-10	-10	+46, -26	+46, -26	25	5
GS 80/09/19	50	-40	140	-27, +107	-27, +107	13	4
GN 80/09/20	45	10	10	+21, -41	+21, -41	11	3
H 80/09/20	40	40	-140	-138, +58	-138, +58	12	3
I 80/10/18	115	70	110	+7, -147	+173, -33	9	3
L 80/10/22	115	70	110	-15, -125	-165, -55	11	3
M 80/10/22	250	40	-40	-177, +97	-3, +83	2	1

INFERRED FLOWS OF ELECTRIC CURRENTS IN SOLAR ACTIVE REGIONS

Y.J. Ding and Q.F. Hong, Yunnan Observatory
Academia Sinica, Kunming, China

M.J. Hagyard and A.C. DeLoach,
Marshall Space Flight Center
Huntsville, Alabama 35812 USA

Abstract. In this paper we explore techniques to identify sources of major current systems in active regions and their channels of flow. Measured photospheric vector magnetic fields together with high resolution white light and H-alpha photographs provide the data base to derive the current systems in the photosphere and chromosphere of a solar active region. Simple mathematical constructions of active region fields and currents are used to interpret these data under the assumptions that the fields in the lower atmosphere (below 200 km) may not be force-free but those in the chromosphere and higher are. The results obtained for the complex active region AR 2372 (April 1980) are: (1). Spots exhibiting significant spiral structure in the penumbral filaments were the source of vertical currents at the photospheric surface. (2). Magnetic neutral lines where the transverse magnetic field was strongly sheared were channels along which a strong current system flowed. (3). The inferred current systems produced a neutral sheet and oppositely-flowing currents in the area of the magnetic delta configuration that was the site of flaring; we postulate that interaction between these two currents produced enhanced turbulence in the neutral sheet leading to particle acceleration in that region. The scenario of flare onset in the neutral sheet with subsequent particle acceleration along the loop systems involved is consistent with the observed flare brightenings observed in H-alpha.

I. Introduction

Electric current systems are presumed to exist in solar active regions, perhaps in the form of the filamentary electric currents proposed by Rabin and Moore (1984) as the source of heating the lower transition region, or as large-scale current systems in solar prominences. Indeed, electric currents are required to flow in the upper atmosphere of flare-producing active regions, providing the superpotential energy that is ultimately released in the form of a solar flare. Observational evidence for current systems is derived from studying the morphology of solar features where magnetic fields are known to play fundamental roles. For example, by studying high-resolution, white-light observations of sunspot groups, scientists from the Yunnan Observatory have inferred the presence of vertical electric currents flowing from spots that exhibit a pronounced spiral structure in the penumbral filaments (Solar Division of Yunnan Observatory, 1974a,b); an example of such a spiral sunspot is shown in Figure 1. From studying a large number of such

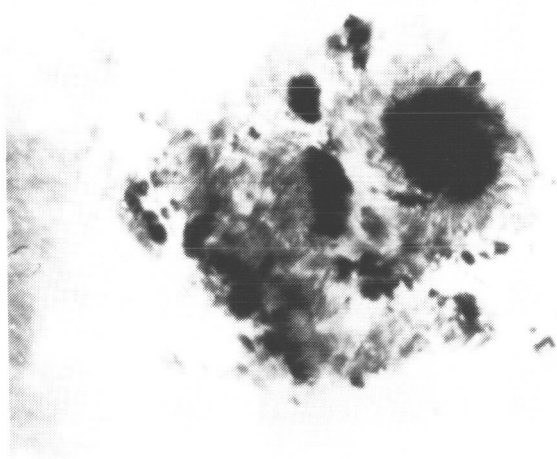


Fig.1

Fig. 1. A large spiral sunspot with positive (north) polarity. This spot was the leader spot in a large sunspot group that was on the disk October 26, 1972. The inferred vertical electric current in it flowed downwards (Ding et al., 1977).

sunspot groups, the Yunnan observers found that there is a close correlation between long-lived spiral sunspots and the occurrence of proton flares.

Direct measurements of electric currents in active regions can be derived from observations of the vector magnetic field using the relation between current and the rotation (curl) of the field, but these measurements give us information only on the vertical component of the current. However, even this limited information has proven valuable since scientists have found correlations between the locations of these vertical currents and various manifestations of solar activity such as flares (Moreton and Severny, 1968; Hagyard et al., 1984) and heating leading to enhanced ultraviolet emission (DeLoach et al., 1984).

In this paper we present results of a study to determine the large-scale current systems flowing in an active region, using observations of the morphology of the photospheric and chromospheric magnetic field together with measurements of the vertical currents derived from vector magnetic field data and some simple mathematical constructions for the magnetic field and resulting current systems. In the following section we discuss first the current systems flowing at the photospheric level in a simple unipolar sunspot, using an analytical model for the magnetic field. In Section 3 we extend the model to the case of a simple bipolar field and discuss the different photospheric current systems that result when the observed transverse field connecting the two polarities has either a 'potential' or a 'sheared' configuration in the vicinity of the magnetic neutral line (where the line-of-sight field between the two polarities vanishes). In Section 4 we indicate how observations of H-alpha fibrils can be used to trace the chromospheric current systems and how these are connected to the photospheric currents through observations of vertical currents at the photosphere. In the last section, we apply these techniques to infer the large-scale currents flowing in a flare-productive active region and relate these current systems to the morphology of the flares that occurred.

II. Electric Currents in Unipolar Spots

To investigate the currents in a simple unipolar sunspot, we used a model of a cylindrically symmetric magnetic field (Allen, 1973) given by the following equations:

$$\begin{aligned} \mathbf{B} &= (B_r, B_\phi, B_z), \\ B_r &= B_z \tan \theta, \\ B_\phi &= 0, \\ B_z &= B_0 \exp(-a \rho^2) f(\rho, z^*), \end{aligned} \tag{1}$$

and $\theta = \theta_0 \rho$, with $\theta_0 = 75^\circ$.

In these expressions, B_z , B_ϕ and B_r are the vertical, azimuthal and radial components of the vector magnetic field \mathbf{B} respectively, θ is the angle of \mathbf{B} to the line-of-sight (i.e., to the normal to the photospheric surface), B_0 is the magnitude of the field at the center of the sunspot at $z = 0$, R is the

outer penumbral radius, ρ is the relative radial coordinate ($\rho = r/R$), and z^* is the relative height ($z^* = z/R$). The function $f(\rho, z^*)$ is determined from the condition that the field have zero divergence:

$$\nabla \cdot \mathbf{B} = 0. \quad (2)$$

This leads to the following differential equation for $f(\rho, z^*)$:

$$\frac{\partial \ln f}{\partial z^*} + \tan \theta \frac{\partial \ln f}{\partial \rho} + (\tan \theta)/\rho + \theta_0/\cos^2 \theta - 2a\rho \tan \theta = 0. \quad (3)$$

The boundary conditions on $f(\rho, z^*)$ are:

$$f(\rho, z^*) \rightarrow 0 \text{ for large } z^*$$

$$\text{and } f(\rho, 0) = 1, \quad (4)$$

$$\text{so } B_z(\rho, 0) = B_0 \exp(-a\rho^2),$$

with the values of $B_z(\rho, 0)$ derived from the observed line-of-sight magnetic field. A particular form of $f(\rho, z^*)$ that satisfies these conditions is

$$f(\rho, z^*) = \exp \left[- \sum_{n=1}^{\infty} (z^*)^n \psi_n(\rho) \right]. \quad (5)$$

Substitution of this expression for $f(\rho, z^*)$ into Equation (3) yields the following equations for ψ_n :

$$\begin{aligned} \psi_1(\rho) &= (\tan \theta)/\rho + \theta_0/\cos^2 \theta - 2a\rho \tan \theta, \\ \psi_2(\rho) &= -(1/2) \tan \theta \psi_1', \\ \psi_n(\rho) &= -(1/n) \tan \theta \psi_{n-1}' \text{ for } n = 2, 3, \dots \end{aligned} \quad (6)$$

Using Ampere's Law,

$$\mathbf{j} = (1/4\pi) \nabla \times \mathbf{B}, \quad (7)$$

we can derive the currents from the field, obtaining

$$J_r = 0,$$

$$J_z = 0,$$

$$J_\phi = [B_z/(4\pi R)] [2a\rho - \tan\theta \psi_1 + z^* (\psi_1' - 2\psi_2 \tan\theta) + (z^*)^2 (\psi_2' - 3\psi_3 \tan\theta) + \dots + (z^*)^n (\psi_n' - (n+1)\psi_{n+1} \tan\theta) + \dots] \quad (8)$$

Thus, for a unipolar sunspot with no azimuthal field component - and thus one where the observed transverse field is radial - only azimuthal currents can occur, and they will flow across the magnetic field. This means that these currents are restricted to the lower chromosphere and photosphere where the magnetic field does not have to be force-free, i.e., below the height of ≈ 200 km where the plasma beta is $\approx 0.22 - 0.61$ and the currents can cross the field. For low heights ($z \lesssim 200$ km) and $R \geq 6$ arcsec, we have $z^* < 0.05$ and the terms involving z^{*2} , z^{*3} , ... can be neglected in the expression for J_ϕ . The radial variation of this approximate form for J_ϕ at different values of z^* is shown in the curves of Figure 2. These curves indicate that the azimuthal current has the same sign as the vertical field B_z everywhere above $z^* = 0.02$, but at low heights, $z^* \leq 0.01$, the current near the outer boundary of the sunspot ($\rho \geq 0.85$) flows in the reverse direction.

To summarize these results for this model of a unipolar sunspot, we find:
(1). There can be no radial or vertical components of the current \underline{J} in a cylindrically-symmetric unipolar sunspot with no azimuthal component of the magnetic field.

(2). The only component of the current \underline{J} is the azimuthal component J_ϕ which flows along the isogauss contours of B_z and across the magnetic field; its direction is in accordance with the right-hand rule. In the outer penumbra, the current J_ϕ reaches its maximum value of $\approx 0.016 \text{ A m}^{-2}$ for $B_z = 300 \text{ G}$ and $R = 10$ arcsec.

(3). The Lorentz force given by $\underline{F} = \underline{J} \times \underline{B}$ has two components, F_r and F_z . Because of the change in direction of the azimuthal current, for $z^* \geq 0.02$ the field lines will be pushed down by the Lorentz forces while, for $z^* \leq 0.01$ and $\rho \geq 0.85$, they will be pushed up; this effect produces a 'trumpet-shaped' configuration for the field lines as indicated in Figure 3.

III. Electric Currents in a Simple Bipolar Region

Observations of single unipolar spots are rare; more often two sunspots with opposite polarities are observed in close proximity, forming a bipolar sunspot group. If the magnetic fields of two sunspots forming a bipole were similar to that of Equation (1), they would produce two parallel currents near the neutral line (the contour $B_z = 0$) and on either side of the neutral line at photospheric levels ($z \leq 200$ km), as indicated in Figure 3. The total horizontal current system flowing between the two spots through the area ABC in Figure 3 is given by

$$I_\phi = \int_{ABC} J_\phi \, dS \quad (9)$$

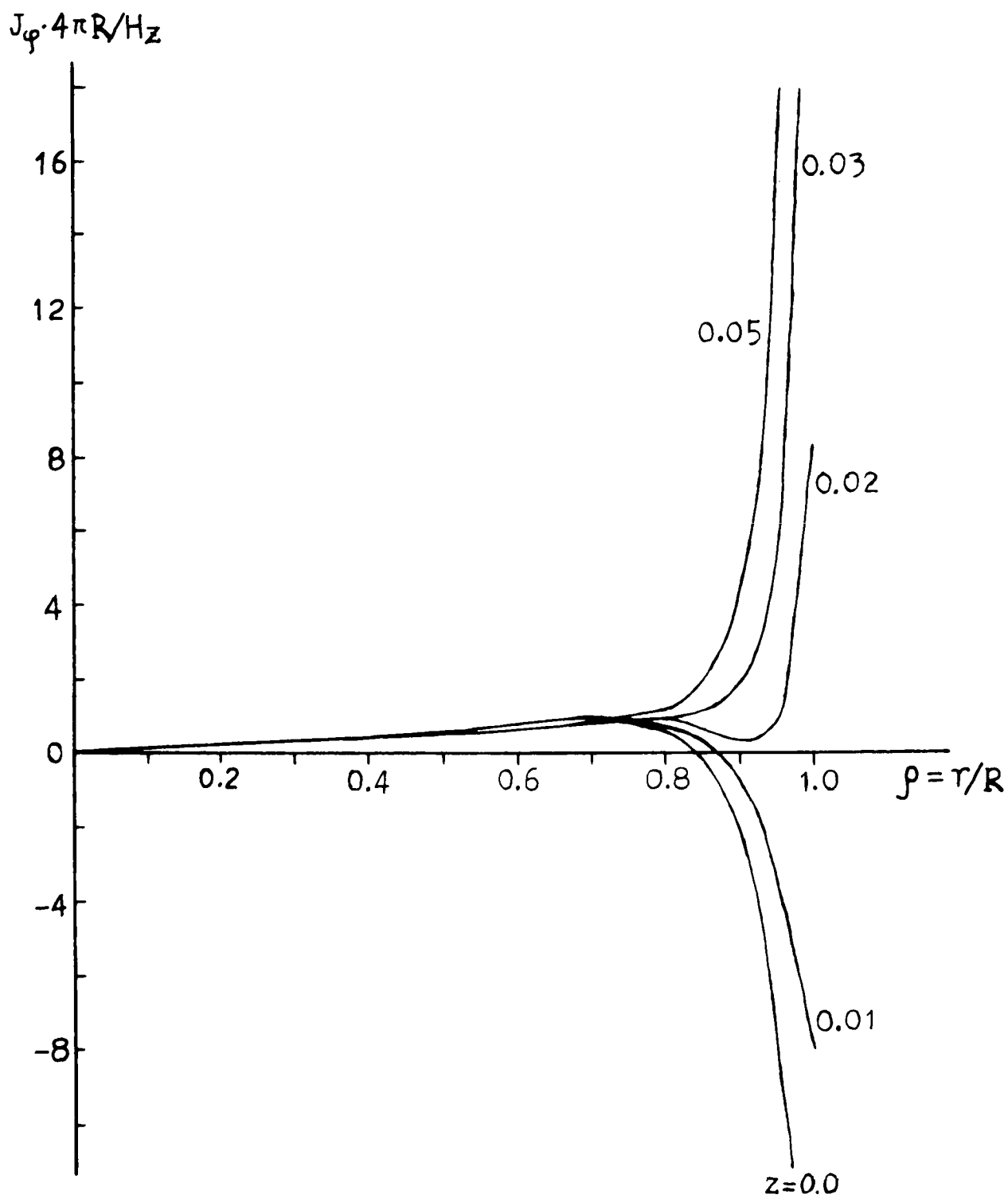


Fig. 2. Variation with radius and height of the azimuthal current density J_ϕ in a simple unipolar sunspot.

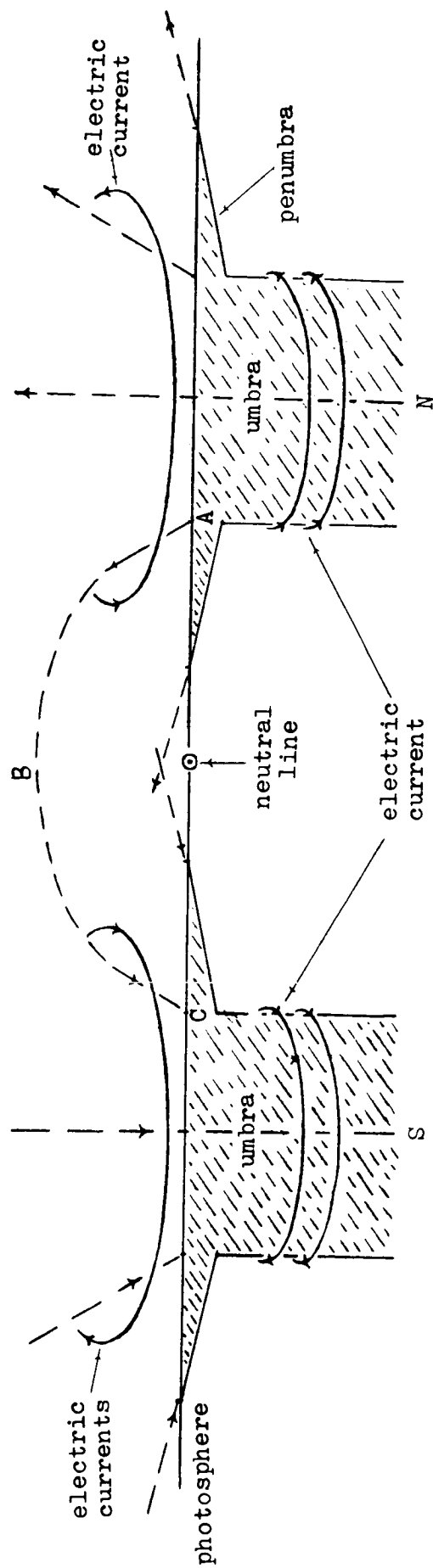


Fig. 3. A sketch of the azimuthal currents in a bipolar sunspot group.

or

$$I_{\phi} = \frac{1}{4\pi} \int_{\widehat{ABC}} \underline{B} \cdot d\underline{l} + \frac{1}{4\pi} \int_{\vec{CA}} \underline{B} \cdot d\underline{l}. \quad (10)$$

The path of integration in these equations is along the arch \widehat{ABC} with A and C representing the points where the line joining the two sunspots' centers crosses the respective umbral boundaries of the two spots. In this construction for the current I_{ϕ} , the first term in Equation (10) involves the magnetic field at heights where it generally cannot be measured, whereas the second integral can be evaluated from observational data of the transverse magnetic field B_T .

If the magnetic field of this bipole in the area ABC can be described by a model similar to that discussed in the previous section which led to the expression given in Equation (8) for the current density J_{ϕ} , then we can make some quantitative statements about the relative magnitudes of the two terms in Equation (10). We found that for small z^* , $J_{\phi}(\rho, z^*) > 0$ except in a small volume where $z^* \leq 0.01$ and $\rho \geq 0.85$; this means that the total integral I_{ϕ} must be greater than or equal to zero by virtue of Equation (9). If the arch \widehat{ABC} is taken along a field line from the positive to the negative polarity, then the first term in Equation (10) will be positive. For a radial field along the path CA, the second term in that equation will be negative since the field \underline{B} and the path increment $d\underline{l}$ are oppositely directed. Therefore, for a nearly radial field, the first term must be larger than the second one when the arch \widehat{ABC} is near the height of 200 km where J_{ϕ} is greater than zero for all ρ and z^* .

In many active regions, however, the field between opposite-polarity spots appears non-radial or sheared in observations of the transverse component (B_T) of the magnetic field or of penumbral filaments seen in white light; this is especially true along the magnetic neutral line where flaring occurs. To underscore this difference, we reconstruct Equation (10) in a manner similar to the technique of Hagyard et al. (1981), and write the total current as the sum of two terms, I_0 and I_{SH} :

$$I_{\phi} = I_0 + I_{SH}, \quad (11)$$

where

$$I_0 = \frac{1}{4\pi} \int_{\widehat{ABC}} \underline{B} \cdot d\underline{l} - \frac{1}{4\pi} \int_{AC} B_T d\underline{l} \quad (12)$$

$$\text{and} \quad I_{SH} = \frac{1}{4\pi} \int_{AC} B_T (1 + \cos \beta) d\underline{l}, \quad (13)$$

where β is the angle between B_T and $d\underline{l}$ along AC. In this formulation the current I_0 is that which would be evaluated for the total current I_{ϕ} if the field along AC appears "potential" or non-sheared, i.e., the field is radial across the neutral line between the two spots. It should be noted that the field is not necessarily potential throughout, but only appears to be potential along AC. For this reason, we call I_0 the "groundstate" current

and note that it cannot be directly evaluated from the observed photospheric field (although from the arguments in the last paragraph, it is positive near the height of 200 km). However, if the observed field along AC appears sheared, then the total current I_ϕ will be the sum of this 'groundstate' current I_0 and the component I_{SH} which can be evaluated from the observed field along the photospheric path AC. Thus this formulation allows us to calculate the currents I_{SH} that are the direct result of the shear in the field independent of the integral over the arch ABC where the field cannot be measured but which contributes only to the 'groundstate' current. This current I_{SH} represents an 'excess' current caused by the sheared transverse field; it flows close to and along the sheared neutral line and in the direction of I_0 . The value of I_{SH} can be estimated as

$$I_{SH} \approx (a \times b) \times 1.2 \times 10^{12} \text{ A} \quad (14)$$

if $B_T = a \times 1000 \text{ G}$, $AC = b \times 20 \text{ arcsec}$ and the angle β is 90° .

IV. Chromospheric Currents and Their Photospheric Source

The above model of a bipolar sunspot group is very simplistic. Generally the magnetic fields observed in active regions are much more complex, particularly in regions where flares occur. In these regions with realistic magnetic structures, there are radial and vertical currents (J_r and J_z) as well as the azimuthal current J_ϕ discussed above. Along the neutral line, we would expect a large excess azimuthal current I_{SH} to exist; such a current system might be connected to the radial currents flowing towards or away from sunspots or to the vertical currents flowing through the photospheric surface. Observations can tell us something about I_{SH} and J_z in the photosphere, but not J_r . Also, the morphology of their connections at this height is obscure since the field is not constrained to be force-free and currents can flow across the magnetic field and penumbral filaments.

The situation is less complex at the chromospheric level above the height of 200 km where the field is force-free so that the currents must be parallel to the magnetic field (Nakagawa et al., 1971; Raadu and Nakagawa, 1971). It is well established that the fibrillar structure seen in H-alpha observations traces the horizontal magnetic field in the chromosphere (Veeder and Zirin, 1971; Foukal, 1971). Since the field there is force-free, the fibrils must trace the horizontal current systems, also. For example, the fibrillar structure of old unipolar sunspots is generally observed to be radial; we interpret this as indicating there are no azimuthal currents flowing in the chromosphere above these sunspots. Conversely, we expect that azimuthal currents will be present only when the transverse field is sheared, that is, only when we observe a vortical structure in the H-alpha fibrils. This means that only the sheared component I_{SH} of the azimuthal current system I_ϕ will be present at chromospheric heights since the I_0 component is constrained to regions where the field is not force-free, i.e., where the current J_ϕ is non-zero in the region of radial magnetic fields. In addition we expect radial and vertical components of currents to flow along the magnetic fields in the chromosphere.

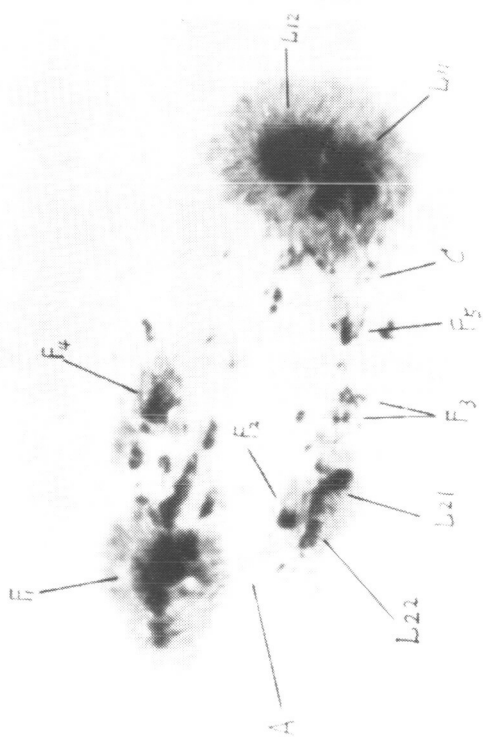
To summarize, the H-alpha fibril structures tell us the following about chromospheric currents:

- (1). Radial H-alpha fibrils trace the systems of radial currents J_r .
- (2). Sheared (relative to the neutral line) fibrils trace the excess current I_{SH} .
- (3). The footpoints of H-alpha fibrils connecting into areas of opposite polarity indicate areas where vertical currents of opposite flows should exist; these areas should be observed in photospheric measurements of J_z unless they are too weak to be detected.

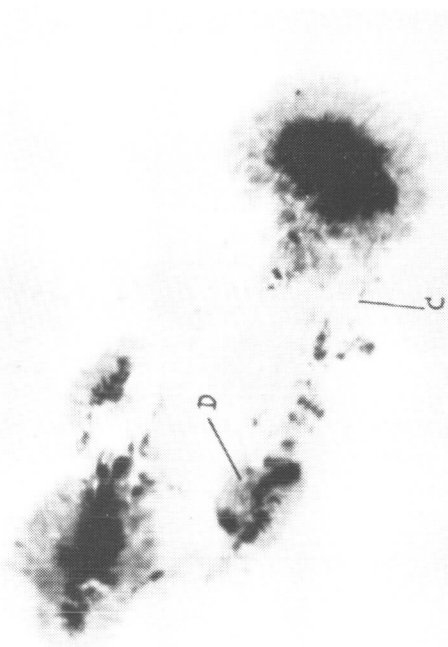
The technique then is to use observations of the transverse magnetic field and H-alpha fibrils and filaments to trace the large-scale current systems of an active region. The transverse magnetic field allows calculations of J_z and I_{SH} , and the sheared filaments along the neutral line trace the connections of areas of $+J_z$ along the path of flow of I_{SH} . This technique will be applied in the following section for an active region that was observed in April 1980.

V. Electric Current Systems in AR 2372

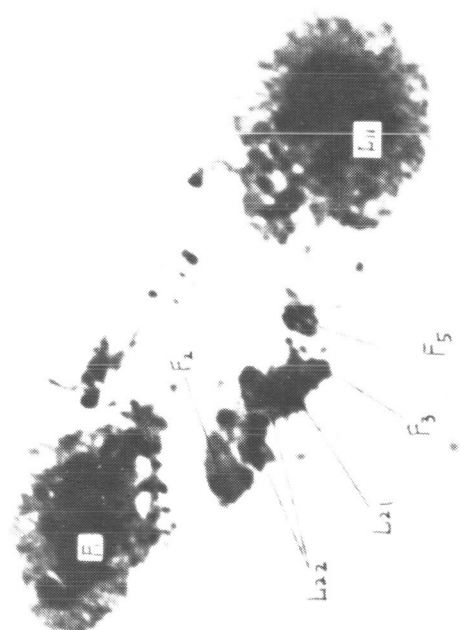
During the period of the Solar Maximum Year the Solar Observatory of the Marshall Space Flight Center operated the MSFC vector magnetograph (Hagyard et al., 1982) on a regular basis, thereby obtaining vector magnetic field data for a large number of active regions, many of which produced significant flares. One of these was AR 2372 which was on the solar disk in the early part of April 1980. The region is shown in Figure 4 as it appeared during the 6th and 7th of April. As discussed by Wu et al. (1984), the field interconnections that were observed in this region on the 5th were realigned as a result of the internal motions of the spots L_{21} and F_2 . Initially L_{21} and F_2 formed a connected bipole (and a magnetic delta configuration) as did the two large sunspots L_{11} and F_1 . Other possible magnetic connections were between the eastern portion of L_{21} (L_{22}) and F_1 . During the period April 5th to the 6th, the positive spot L_{21} moved westward sweeping up negative fields (F_3 , F_5) ahead of it and forming a new magnetic delta configuration (L_{21} , F_3). At the same time the negative spot F_2 which was originally to the north of L_{21} moved eastward toward the large negative spot F_1 . The resulting field realignments produced connections between the negative fields F_3 , F_5 , etc., that were built up and the leader spot L_{11} , and the original connections between L_{21} and F_2 were strongly sheared. In addition the motion of F_2 probably produced a shear in the interconnection between L_{22} and F_1 . In Figure 5a the vector magnetic field of this region on the 6th is shown and reveals the sheared nature of the interconnections between L_{21} and F_2 and L_{22} and F_1 . This is especially evident when the field is compared with a potential field, shown in Figure 5b, that was derived using the line-of-sight field of Figure 5a. In Figure 6 the vertical electric currents derived by Krall et al. (1982) from the vector magnetic field data are shown for the areas of the sunspots F_2 , L_{22} , L_{21} , F_3 and F_5 . These spots are also identified in the H-alpha photograph (from Big Bear Solar Observatory) shown in Figure 7. The set of data shown in Figures 4-7 has been used in conjunction with the analytical constructions developed previously to derive the current systems for the 6th of April.



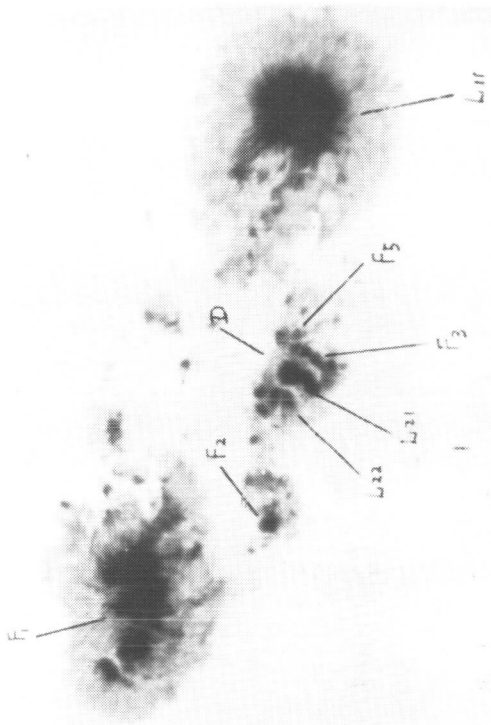
a.6/0135UT



b.6/0335UT



c.6/1940UT (Big Bear Obs.)



d.7/0530UT

Fig. 4. The sunspot group forming AR 2372 on April 6 and 7, 1980. (a). The sunspots' grouping on April 6 at 01:35 UT. (b). The active region at 03:35 UT on the 6th. (c). Big Bear filtergram at 19:40 UT on the 6th (courtesy of Dr. H. Zirin). (d). The region early on the 7th at 05:30 UT.

Fig.5a

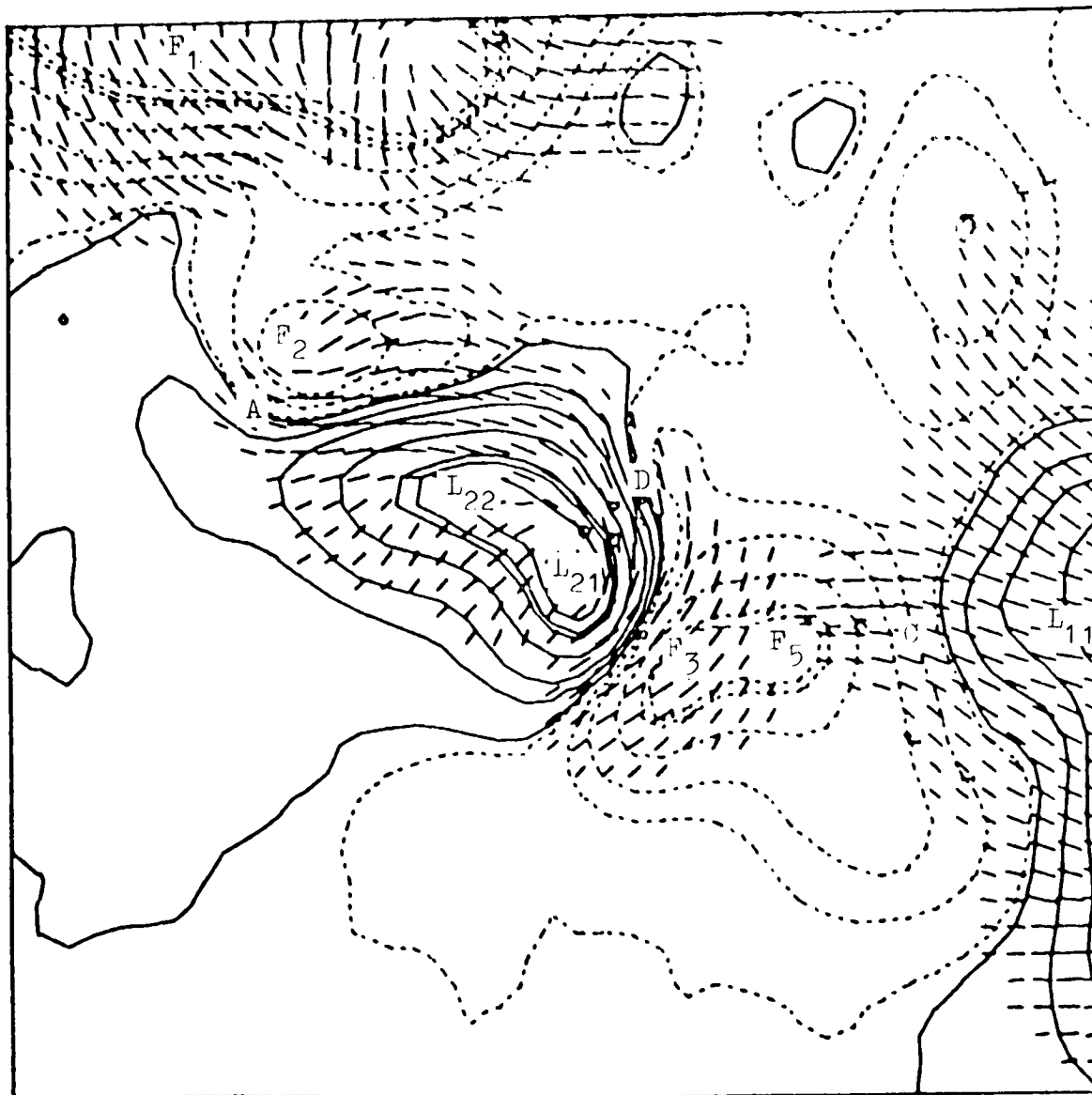
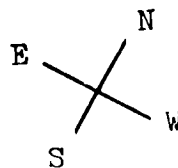


Fig. 5. The vector magnetic field in the area of the magnetic delta configuration of AR 2372 on April 6. The contours represent the line-of-sight component of the magnetic field with solid (dashed) contours depicting positive (negative) fields. The line segments represent in length and direction the magnitude and orientation of the transverse component of the magnetic field. (a). The observed field at 21:00 UT on the 6th. (b). A potential field calculated from the line-of-sight field of (a).

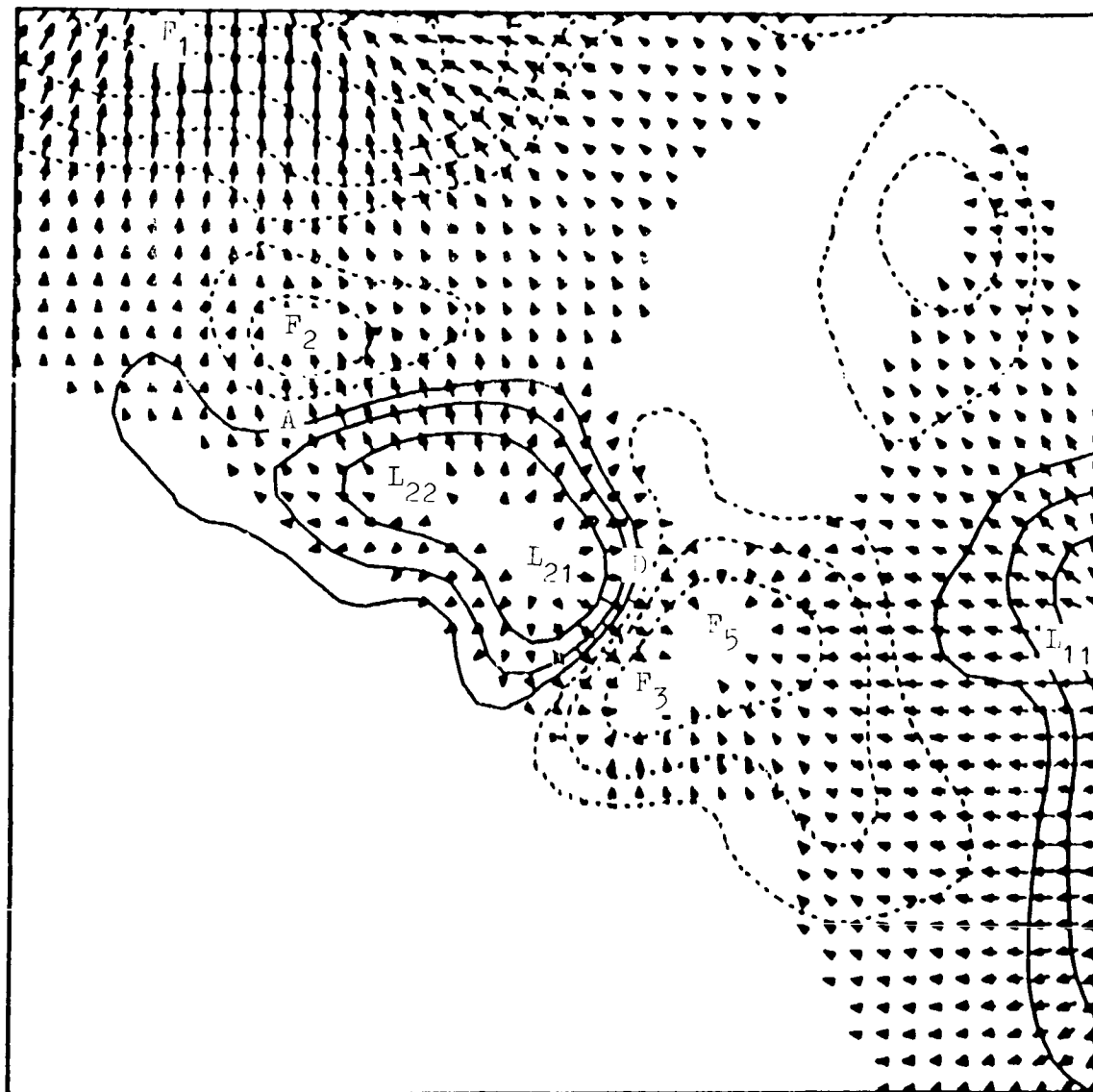
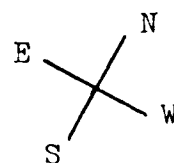


Fig.5b



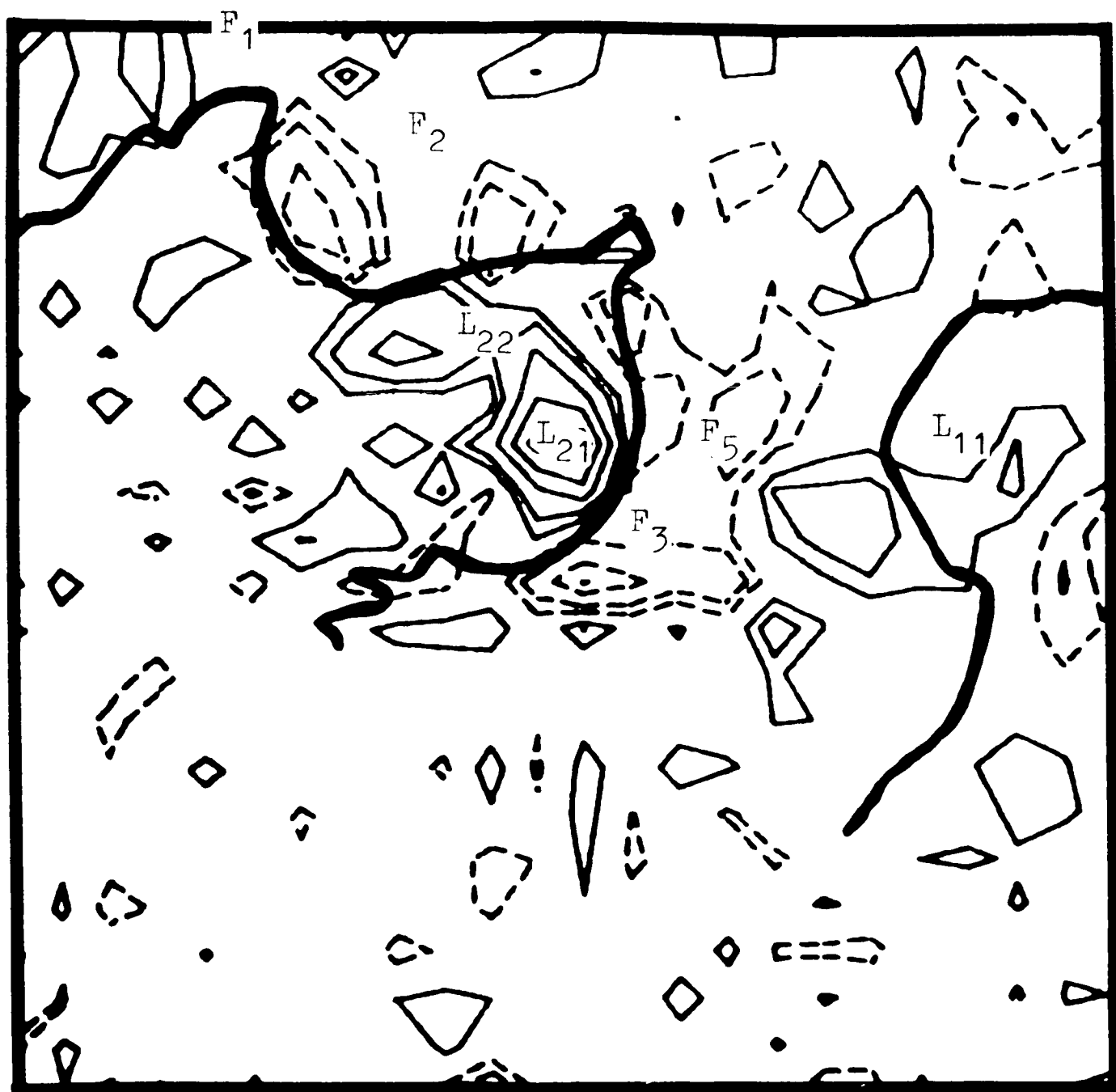


Fig. 6a

Fig. 6. The distribution of the vertical electric current densities in the area of the magnetic delta configuration on April 6, 1980 (from Krall et al., 1982). Solid (dashed) contours represent currents flowing out from (into) the photosphere. The heavy solid line represents the magnetic neutral line to aid in orientation with the magnetic field shown in Figure 5a. (a). The vertical current densities at 19:08 UT. (b). The vertical currents at 21:00 UT.

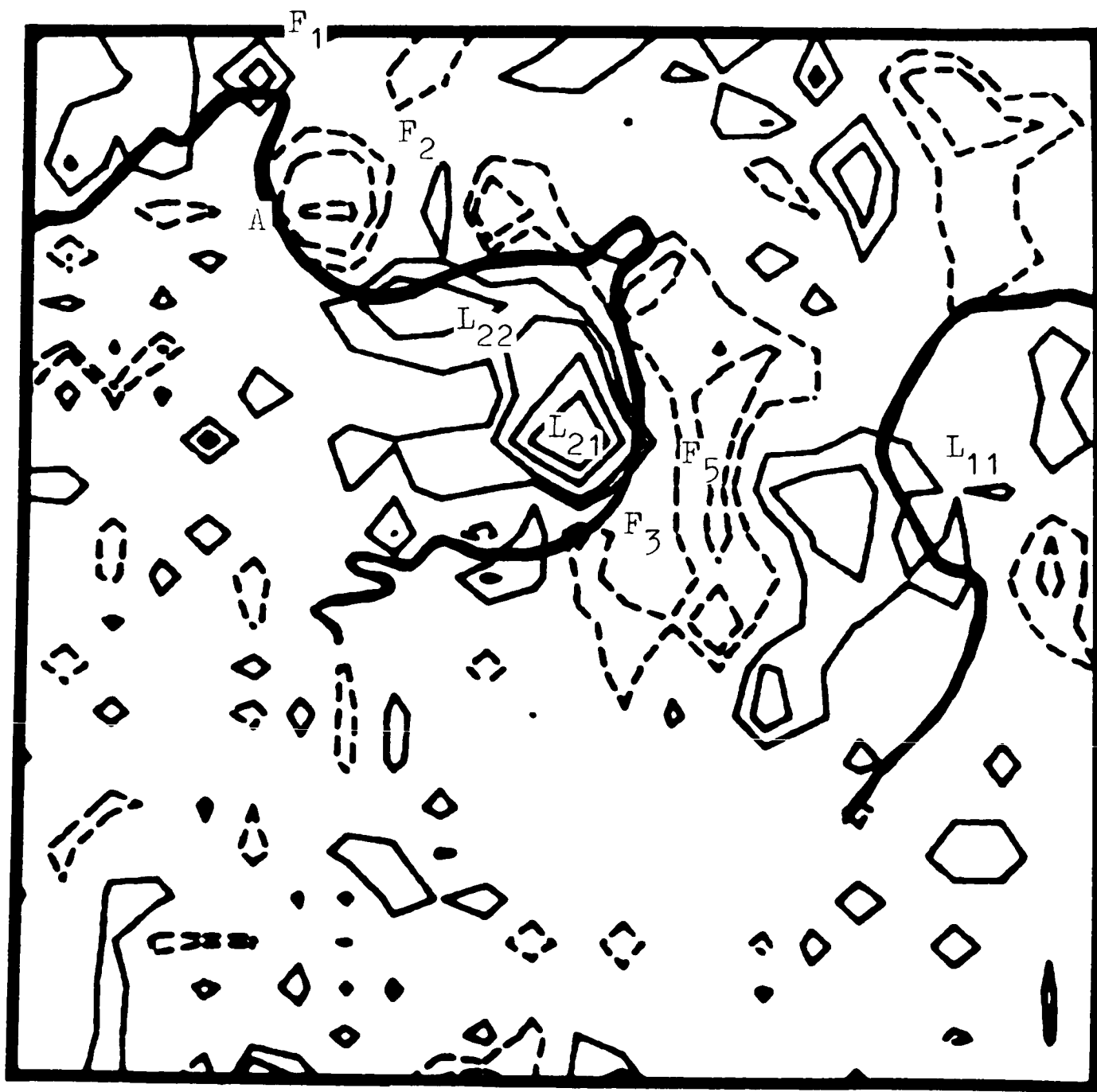


Fig. 6b

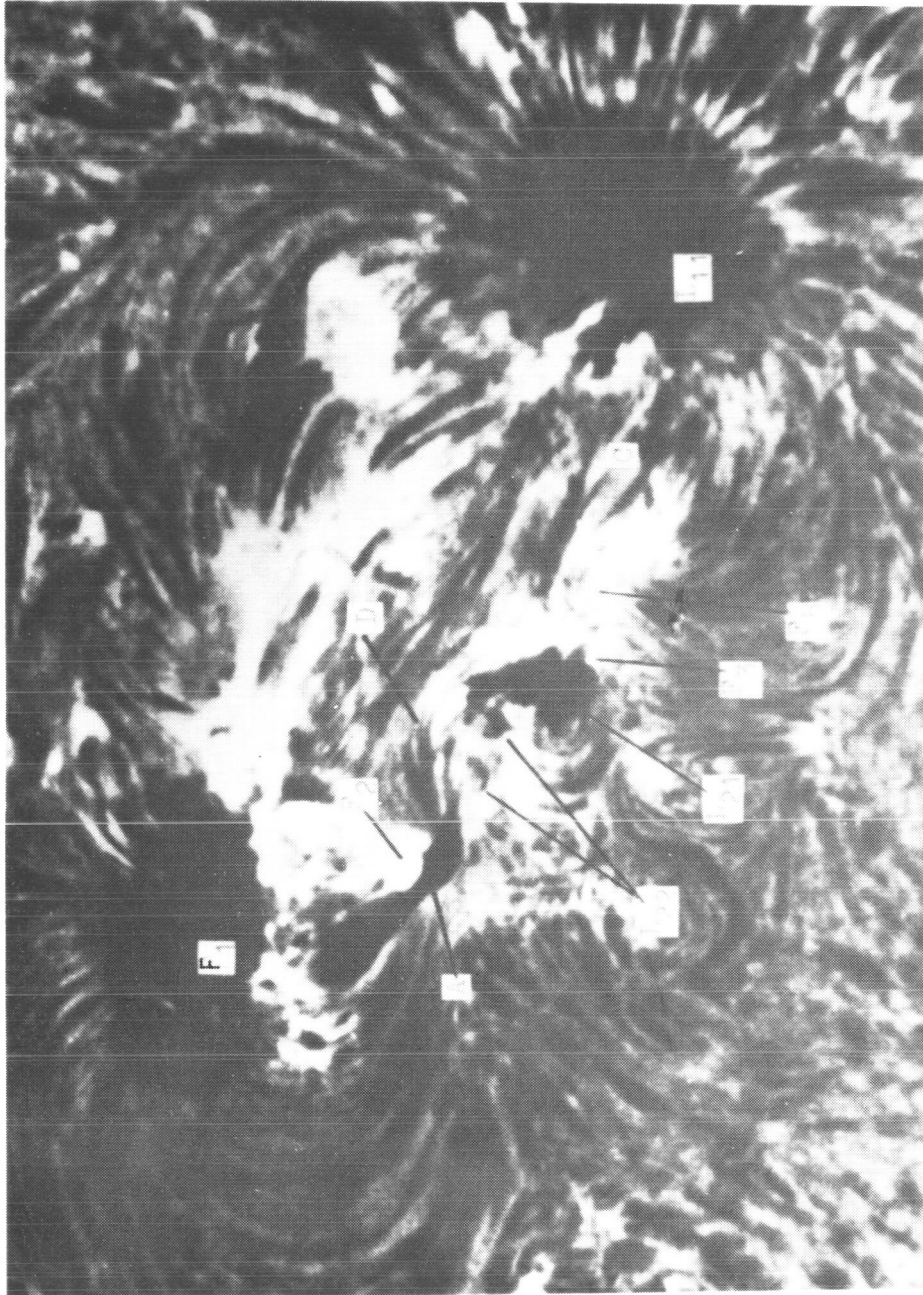


Fig. 7. Centerline H-alpha filtergram of AR 2372 on April 6, 1980 at 19:39 UT. (From Big Bear Solar Observatory, courtesy of Dr. H. Zirin.)

(1). A current system from L_{21} to F_2 along a sheared neutral line. The spiral sunspot L_{21} with a predominantly counterclockwise penumbral pattern was located at the same position as a counterclockwise vortex in the observed transverse magnetic field (Figure 5a) and a positive J_z concentration (Figure 6). This current direction is consistent with the direction of the observed spiral pattern. By estimating the average current density and the area from Figure 6a, the current strength I_z in L_{21} was calculated to be about 1×10^{12} A. In Figure 5a a segment of the magnetic neutral line is indicated by the letter D. Along this segment of the neutral line the sheared transverse field was very similar to the sheared penumbral filaments (D) which were seen to persist for 28 hours in Figure 4a,b and d. Deviating as they did from the configuration of a potential field, they suggest that a strong excess current I_{SH} was flowing along the segment D and eastward. By taking $a = 0.7$, $b = 0.5$ and $\theta = 90^\circ$ in Equation (14), the current I_{SH} flowing in the lower atmosphere along the neutral line is estimated to be 0.4×10^{12} A. In Figure 7 the same area as seen in H-alpha is again designated by the letter D. This area shows some dark fibrils connecting L_{21} with F_2 , from which we infer that a current system was flowing in the chromosphere along the fibrils D from L_{21} ($J_z > 0$) to F_2 ($J_z < 0$). Since both the photospheric and chromospheric currents flowed in the same direction along the neutral line, we infer that the two currents belonged to one current system that was gradually rising upwards. The estimated current flowing into F_2 was 0.3×10^{12} A (calculated from the average J_z and area of F_2 in Figure 6a). This is approximately equal to the current estimated for I_{SH} but somewhat less than that estimated to emerge from L_{21} .

(2). A current system flowing from L_{22} to F_1 along a sheared neutral line.

Sheared penumbral filaments and a sheared transverse magnetic field were observed along the segment "A" of the magnetic neutral line indicated in Figures 4a and 5a. These data again suggest an excess current I_{SH} flowed eastward along the sheared neutral line A from L_{22} to F_1 . Using Equation (14) the current I_{SH} along A can be calculated; it is less than that along D since the transverse field strength in this area was weaker. A positive vertical current density J_z was calculated in the area of the spot L_{22} ; from its average value and the area of L_{22} we estimate a positive current of 5×10^{11} A flowing upwards from L_{22} . Because a negative vertical current appears to exist in the area of F_1 , we suggest that a photospheric current flowed from L_{22} to F_1 along the neutral line A. The H-alpha data show a dark sunspot filament lying along A connecting F_1 and L_{22} . Thus we again infer that the photospheric and chromospheric currents formed a single current system that was rising upwards.

(3). A radial current flowed from L_{11} to F_3 , F_5 .

Some dark penumbral-like filaments were observed in the area "C" indicated in Figure 4a (01:35 UT on the 6th). From this observation we infer that the fibrous magnetic field linking the bipolar group L_{11} and F_3 , F_5 , etc., had emerged from the convection zone into the photosphere prior to 01:35 UT on the 6th. In the H-alpha filtergram we see several long, radial fibrils connecting L_{11} with F_3 and F_5 . Calculated vertical current densities show positive currents in L_{11} and negative ones in F_3 and F_5 . Thus we

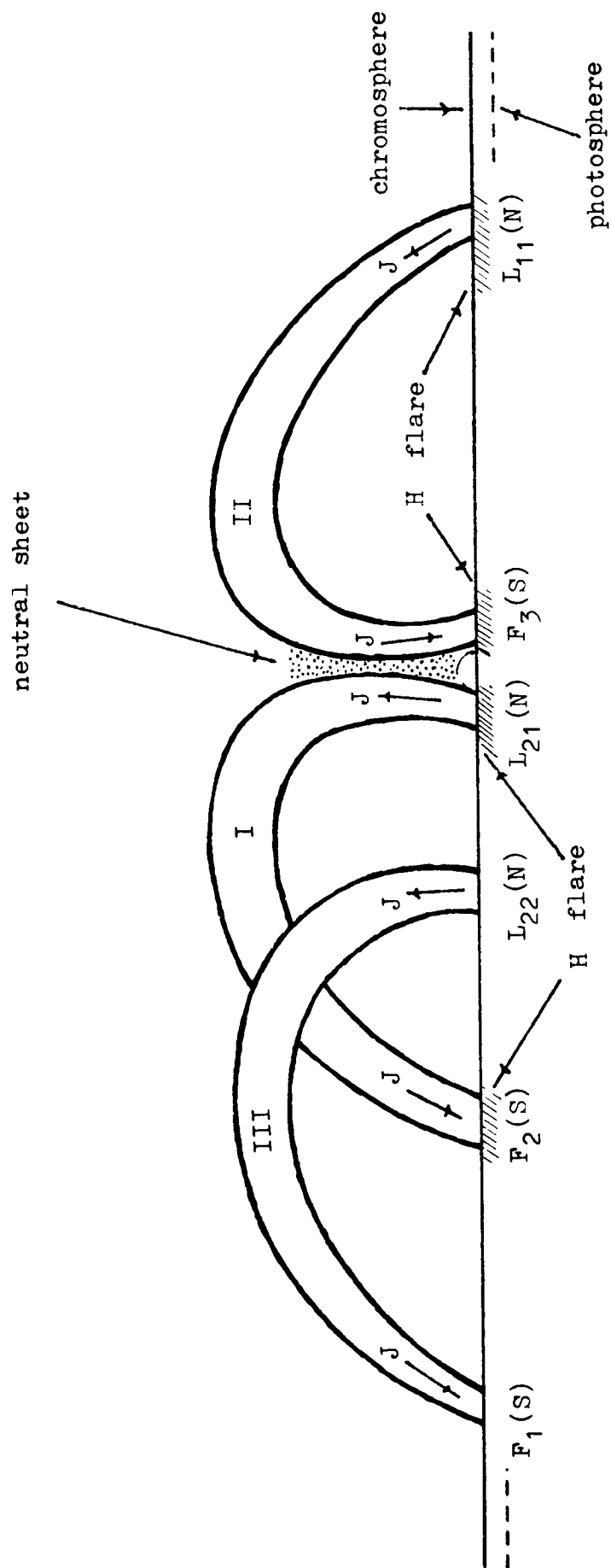


Fig. 8. A sketch of the three major current and magnetic loop systems inferred from observational data. Locations of H-alpha flare brightenings are indicated at the footpoints of the loops.

infer a radial current system flowing along the fibrils from L_{11} to F_3 and F_5 .

In Figure 8 we show a sketch of these three major current and magnetic loop systems which we have derived from the observational data. The flaring activity in this region during the period April 5-7 was concentrated in the area of the new delta configuration (L_{21} and F_3) that was formed by the emergence of loop II and the westward motion of the footpoint L_{21} of loop I.

In addition to the magnetic delta configuration, complexity in this area of the neutral line was enhanced by the oppositely-directed currents and magnetic fields in the footpoints L_{21} and F_3 of loops I and II respectively. The relationship of these complex configurations to the flaring that occurred is interpreted as follows. Superpotential magnetic energy was stored in this area both in the sheared magnetic field and in the neutral sheet between loops I and II. The opposing currents in close proximity in I and II intensified the turbulence in the neutral sheet, accelerating charged particles to high energy and initiating rapid magnetic reconnection. The accelerated particles were projected from the acceleration region in the neutral sheet along the field lines to the chromospheric footpoints where heating occurred as seen in H-alpha. These areas are indicated in Figure 8 by the hatched areas. Thus this current/field system I, II and III explains the observed flare brightenings as seen in H-alpha.

It is interesting to compare these derived loop systems with the four magnetic structures inferred by Machado et al. (1983) from X-ray observations of a flare in AR 2372 at 03:03 UT on April 8. Referring to their Figure 7, the loop designated as B in that figure can be identified with our loop II; their loop C corresponds to either loop I or III in our interpretation. Thus the higher X-ray loops partially fit our derived current system. However, we do not indicate a loop system corresponding to A in the Machado et al. figure that crosses over the highly sheared neutral line between L_{21} and F_3 . Our interpretation was based on the very sheared nature of the transverse magnetic field in this area which indicated that the field was not connecting across the neutral line between L_{21} and F_3 . Since such a connection was inferred from the X-ray data, we conclude that the lower photospheric and chromospheric fields were highly sheared whereas those in the corona were not in the area of the magnetic delta.

Acknowledgement

The authors are extremely grateful to Dr. H. Zirin of the Big Bear Solar Observatory for making available to us the high-resolution H-alpha filtergrams.

References

- Allen, C. W.: 1973, *Astrophysical Quantities*, The Athlone Press.
- DeLoach, A. C., Hagyard, M. J., Rabin, D., Moore, R. L., Smith, J. B. Jr., West, E. A., and Tandberg-Hanssen, E.: 1984, *Solar Phys.* 91, 235.
- Ding, Y. J., Li, W. B., Hong, Q. F., and Li, Z. K.: 1977, *Acta Astronomica Sinica* 18, 39.
- Foukal, P.: 1971, *Solar Phys.* 20, 298.

- Hagyard, M. J., Low, B. C., and Tandberg-Hanssen, E.: 1981, Solar Phys. 73, 257.
- Hagyard, M. J., Cumings, N. P., West, E. A., and Smith, J. E.: 1982, Solar Phys. 80, 33.
- Hagyard, M. J., West, E. A., and Smith, J. B. Jr.: 1984, in Proceedings of the Kunming Workshop on Solar Physics and Interplanetary Travelling Phenomena, Kunming, People's Republic of China.
- Krall, K. R., Smith, J. B. Jr., Hagyard, M. J., West, E. A., and Cumings, N. P.: 1982, Solar Phys. 79, 59.
- Machado, M. E., Somov, B. V., Rovira, M. G., and De Jager, C.: 1983, Solar Phys. 85, 157.
- Moreton, G. E. and Severny, A. B.: 1968, Solar Phys. 3, 282.
- Nakagawa, Y., Raadu, M. A., Billings, D. E., and McNamara, D.: 1971, Solar Phys. 19, 72.
- Raadu, M. A. and Nakagawa, Y.: 1971, Solar Phys. 20, 64.
- Rabin, D. and Moore, R.: 1984, Astrophys. J., in press.
- Solar Division of Yunnan Observatory: 1974a, Acta Astronomica Sinica 15, 25.
- Solar Division of Yunnan Observatory: 1974b, Acta Astronomica Sinica 15, 173.
- Veeder, G. J. and Zirin, H.: 1970, Solar Phys. 12, 39.
- Wu, S. T., Hu, Y. Q., Krall, K. R., Hagyard, M. J., and Smith, J. B. Jr.: 1984, Solar Phys. 90, 117.

COMPARISONS OF SIMULTANEOUS VECTOR MAGNETOGRAMS

M. Makita,* K. Nishi,* M. Shimizu*, S. Hamana,* T. Sakura,** V. M. Grigoryev,***
G. V. Kuklin,*** and V. L. Selivanov***

**Tokyo Astronomical Observatory, University of Tokyo, Mitaka, Tokyo 181 Japan*

***Department of Astronomy, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan*

****Sib-IZMIRAN, Irkutsk 664033, USSR*

Abstract: Simultaneous vector magnetograms were obtained with the different magnetographs of the Okayama Astrophysical Observatory and the Sayan Mountain Observatory in October 1983. The data obtained show a good correspondence in general. The comparison makes clear something on the measuring accuracy of each magnetograph.

1. Introduction

Whether vector magnetograms obtained with different magnetographs are the same or not will be important to see the reliability of the observation and to know the accuracy of each magnetograph. Since the geographical longitudes of the Okayama Astrophysical Observatory, Tokyo Astronomical Observatory, and the Sayan Mountain Observatory, Sib-IZMIRAN, are $+134^\circ$ and $+101^\circ$, respectively, we planned a simultaneous observation program from September 21 to October 20, 1983.

2. Observation

We observed the sunspot group 260 in Russian Solar Data using the Zeeman sensitive line, FeI 5250A, and with the two different types of instruments. Some characteristics are summarized in the following (cf. Stepanov et al., 1975; Makita et al., 1984).

	Okayama Astrophys. Obs.	Sayan Mountain Obs.
Instrument		
Telescope	equatorial	horizontal (coelostat)
Polarimeter	rotating wave plate	two KDPs
Observational Parameters		
Integration time	1 sec	1 sec
Scanning step	10"	4"
Covered area of entrance slit	12" x 0.6"	4" x 2"
Covered wavelength of exit slit (referred to line center)	27 mÅ-80 mÅ	37 mÅ-95 mÅ
Successful Observations		
October 5	3h3m-4h33m UT	2h22m-3h28m UT 4h17m-5h17m UT
October 6	2h57m-4h11m UT	2h26m-3h40m UT

The observing area of Okayama's observation was roughly four times larger. Therefore, its observing period should be shrunk when a more detailed comparison of the time is necessary.

3. Comparison of the Magnetograms

The comparison of the observed data are shown in Figure 1. A quick look shows good correspondence of the circular and linear polarizations, and, in and near sunspots, of the velocity field.

The wing intensity is drawn at 90 and 50 percent levels referred to the photospheric intensity. This is affected by clouds easily, although the polarizations are normalized by the intensity and are not affected by them. Zigzag patterns in the Okayama data, which are also seen in the other lower diagrams, are due to the accuracy of the image guider, since the observation was made by scanning the image in the east-west direction back and forth. This seems to be dependent on the seeing and may be corrected after the observation. By adjusting the guider, otherwise, the Okayama data may have smooth patterns at the cost of the longer observing time.

The longitudinal field of the Sayan data is shown by the steps of 10, 50, 100, 500, and 1000 G. The Okayama data show the degree of circular polarization of 4.0×10^{-3} for the outermost contour and it increases to the inner contours in a geometrical progression of factor 2. The transverse field of the Sayan data shows 200, 500, and 1000 G. Roughly 100 G field prevails outside these contours and seems to show the noise level. The degree of linear polarization of the Okayama data is 2.5×10^{-3} for the outermost contour and increases to the inner contours in a geometrical progression of factor 2. The similar pattern of the polarizations enables us to connect them as in Figure 2. A calibration of the Okayama data is thus made roughly.

The azimuths of the magnetic field agree fairly well in most areas, but give a remarkable discrepancy in the middle of the following sunspot in every case. Since the exit slits are placed in the line wings (see the table), this may not be due simply to Faraday rotation. H alpha filtergrams of the Norikura Solar Observatory and the Big Bear Solar Observatory seem to coincide better with the Okayama's azimuth. Some problem in the Doppler compensator or the different spatial resolution might be responsible for this discrepancy.

The radial velocity of the Sayan data is 0.5 and 1 km/s. The Okayama data show 0.1, 0.5, and 1 km/s. Neglecting the guiding error, the two data are similar in and near sunspots, the amplitude of Okayama data being smaller according to the low spatial resolution. The five minute oscillation will dominate outside the sunspot area and no agreement is expected there. The low amplitude in the second Sayan observation might show a problem with the Doppler compensator when the intensity becomes weak (by clouds in this case).

Acknowledgments: Thanks are due to Prof. H. Zirin of the Big Bear Solar Observatory and Mr. Y. Nakagomi for providing the H alpha pictures on the observing days.

References

- Makita, M., Hamana, S., and Nishi, K., 1984, in this Proceedings.
Stepanov, V. E., Grigoryev, V. M., Kobanov, N. I., and Osak, B. F., 1975, *Issledovaniya Geomag. Aeronomy and Solar Phys.*, 37, 147.

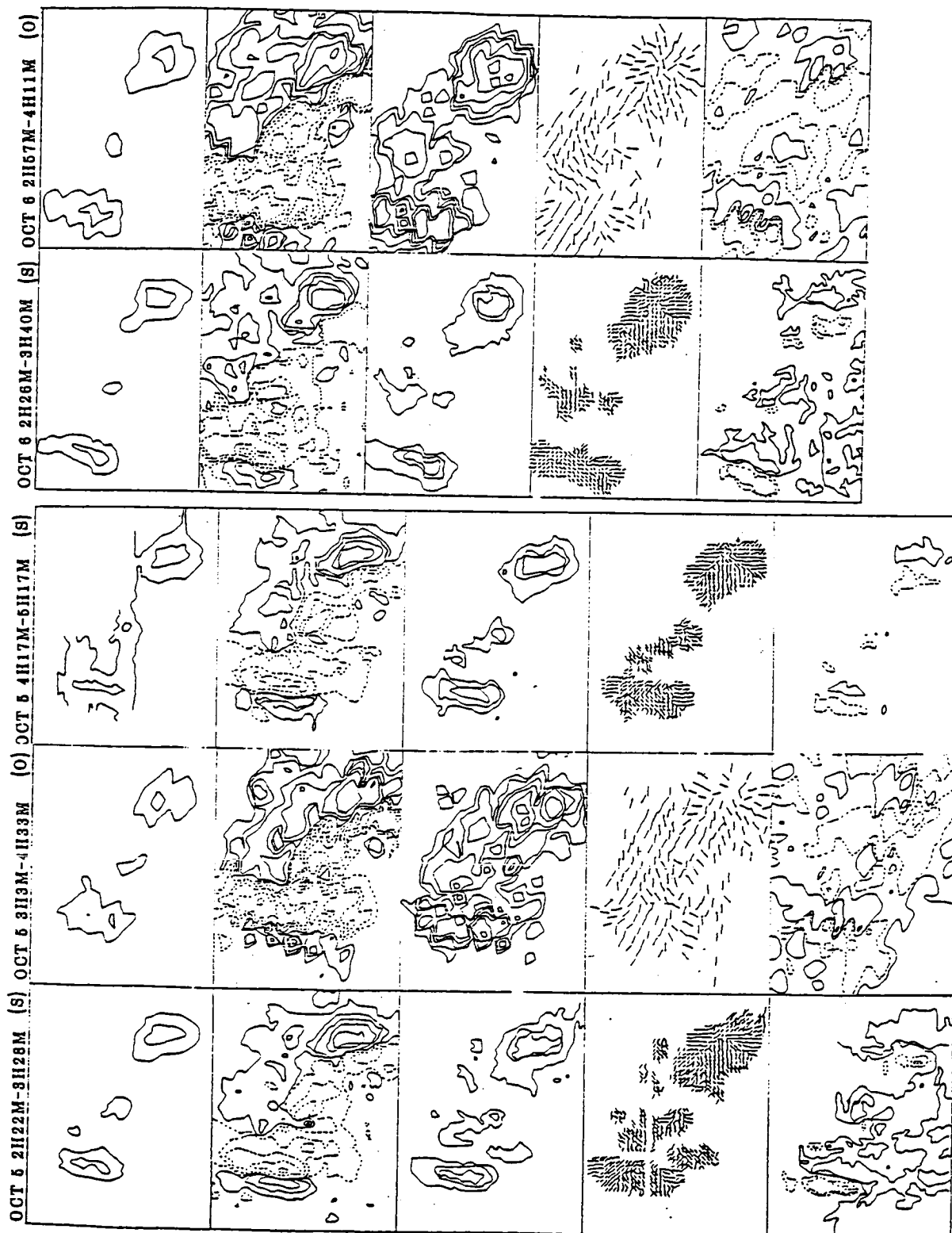


Fig. 1. Comparison of the magnetograms (S: Sayan, O: Okayama). From top to bottom intensity, longitudinal and transverse magnetic fields, azimuth of the magnetic field, and radial velocity. Vertical is parallel to the sun's meridian (up: north, left: east).

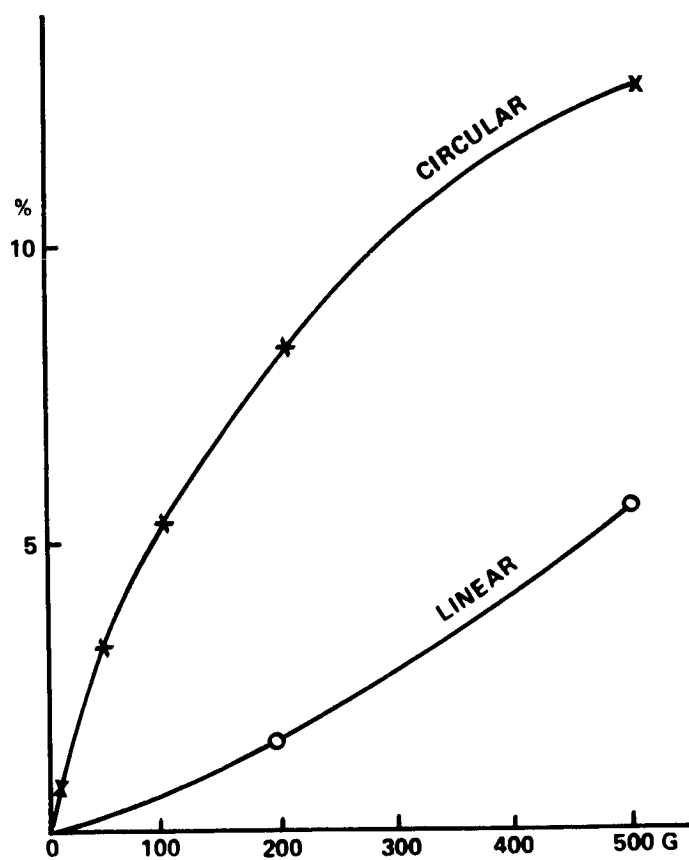


Fig. 2. Relation of the polarization data. Ordinates: the degree of polarization of the Okayama data. Abscissa: magnetic field strength of the Sayan data.

EPHEMERAL REGIONS VS. PSEUDO EPHEMERAL REGIONS

S. F. Martin
Big Bear Solar Observatory
California Institute of Technology

S. H. B. Livi*

J. Wang and Z. Shi**

* Visiting Astronomer from Instituto de Fisica, Universidade Federal do Rio Grande do Sul, 90,000 Porto Alegre, Brazil

** Visiting Associates from Beijing Observatory, The Academy of Sciences, Beijing, China

ABSTRACT

Our new studies of the quiet sun reveal that ephemeral active regions constitute a minority rather than a majority of all the short-lived, small-scale bipolar features on the sun. We retain the definition of an ephemeral region as a feature which appears to originate as a compact bipole and grows as a unit in total flux for at least a short time after its birth. In contrast to the recognized patterns of growth and decay of ephemeral regions, we illustrate various examples of the creation of other temporary bipoles nicknamed "pseudo ephemeral regions." We show that the pseudo ephemeral regions are the consequence of combinations of small scale dynamic processes of the quiet sun including: (1) fragmentation of network magnetic fields, (2) the separation of opposite polarity halves of ephemeral regions as they grow and evolve, and (3) the coalescence of weak network or intra-network magnetic fields. In long-exposure videomagnetograms (~1 min.), having spatial resolution of 2-5 arc seconds, the pseudo ephemeral regions outnumber the real ephemeral regions by about a factor of 2. These new observations offer the possibility of resolving the discrepancies that have arisen in the association of ephemeral regions with X-ray bright points. We suggest that many X-ray bright points may be related to those pseudo ephemeral regions which have begun to exhibit magnetic flux loss. We also suggest that vector magnetograms should also

reveal distinct differences between real and pseudo ephemeral regions similar to the differences that are sometimes seen in H α filtergrams. H α images show the opposite polarities of some ephemeral regions to be connected by fibrils or arch filaments. In contrast, the pseudo ephemeral regions sometimes reveal a fibril, like a small filament, dividing the opposite polarity fields, but no well-defined fibrils connecting the opposite polarities.

I. INTRODUCTION

Ephemeral regions are small, short-lived bipolar magnetic fields. They were recognized and characterized by their short lifetimes in CaII filtergrams years before their magnetic nature became known (Harvey and Martin 1973). Their name was adopted from the tabulations of active regions in Solar Geophysical Data (SGD); short-lived small CaII plages listed for only one or two days are designated as "ephemeral" (for example, see SGD 1963 to 1965). CaII plages are now more frequently called by the more general name "active regions" and hence small Ca plages are now known as "ephemeral active regions" or just "ephemeral regions."

In this paper we clarify the currently recognized properties of ephemeral regions as seen in magnetograms and H α filtergrams. The H α properties of these regions are especially important because H α observations are currently our only direct source of information on the magnetic field component of ephemeral regions parallel to the solar surface; current-day vector magnetographs are not sufficiently sensitive to detect most ephemeral regions.

We further clarify the properties of ephemeral regions by comparing them with other features which can look like ephemeral regions but do not evolve like ephemeral regions. We call these "pseudo" ephemeral regions.

II. THE DATA

We have previously shown that long exposure videomagnetograms from Big Bear Solar Observatory are effective for studying ephemeral regions and other magnetic fields on the quiet sun (Martin 1983). Exposure of a videomagnetogram consists of many successive scans of the TV camera used as the primary detector in the videomagnetograph. The quiet sun images illustrated in this paper consist of 1024 or 2048 scans, which respectively require scan times of 68 or 138 sec. The final digitized image may be recorded both on magnetic tape and as a photographic image of the magnetogram from the initial display on a television monitor. Magnetograms in both of these formats are illustrated in this paper.

III. EXAMPLES OF EPHEMERAL REGIONS AND PSEUDO EPHEMERAL REGIONS

The properties by which we define ephemeral regions are illustrated by an example in Figure 1. The figure shows small sections of a single field of view taken during an 8 hour observing day on 4 September 1983. The oval in the second frame, 1930 (UT) in the left column of images, encloses a new ephemeral region not seen in the earlier image at 1751. The periphery of the positive pole is white and the periphery of the negative pole is black. The grey contour within the negative pole is not a polarity reversal. It is a contour intentionally included to show that the magnetic signal has reached saturation. By the next frame, 2040 UT,

the bipole has increased in total flux as seen by the appearance of additional saturation contours within each pole. Saturation contours of negative polarity are grey and saturated contours of positive polarity are white. This example illustrates the three defining properties of ephemeral regions:

- (1) a new bipole with opposite polarity fields adjacent to each other
- (2) growth of the bipole
- (3) separation of the maxima of the opposite polarity fields from each other

A useful, non-defining property that is often but not always seen is:

- (4) continued separation of the opposite poles

Note that the definition of an ephemeral region can include:

- (5) emergence of one or both poles within pre-existing network
- (6) a lack of exact simultaneity in the appearance of the opposite polarities (assumed to be an effect of field geometry or limited spatial resolution)

In contrast to ephemeral regions, next we illustrate features which temporarily look like ephemeral regions but lack the above defining properties of ephemeral regions.

The first example of such a pseudo ephemeral region is marked by a rectangle in Figure 1. In the first two frames at 1751 and 1930, the bipole looks exactly like the ephemeral region enclosed in the oval at 1930 UT. However, as we follow its evolution in the subsequent frames, we see that it is a disappearing bipole. The flux in both polarities is rapidly reducing relative to neighboring fields. We categorize this feature as a "pseudo ephemeral region," an apparent bipole that does not exhibit any of the defining properties of ephemeral regions except that it appears temporarily to resemble an ephemeral region. This type of pseudo ephemeral region reveals new, important distinctions from real ephemeral regions: (1) the opposite polarity fields move together, (2) they show a mutual loss in magnetic flux (cancellation), and (3) the gradient of the magnetic field between the centers of opposite polarity increases with time and typically exceeds the intra-pole gradient of ephemeral regions.

Small-scale, adjacent opposite polarity magnetic features are very common on the quiet sun. The corner of the rectangle to the lower right of the ephemeral region (Fig. 1) encloses a fragment of positive polarity field adjacent to network of negative polarity. It does not look like an ephemeral region because of the very large imbalance of flux in the

adjacent opposite polarities. However, even in this example, if our only available image were the one at 1751, we could not be certain that the positive polarity was not one of the halves of an ephemeral region which had its opposite pole buried in the negative polarity network. However, the next several frames show that this is not an ephemeral region. The first sign that this feature is not a new ephemeral region is the steady loss of flux in the positive pole. Concurrently there appears to be loss of flux in the negative polarity indicated by the indentation of the negative flux at the point of contact with the small positive fragment of magnetic field. The positive fragment has only a small point left at 2150 and has completely disappeared by 2234.

During a day's observation of the quiet sun, we typically observe many such examples of the mutual loss of flux in closely-spaced opposite polarity fragments of magnetic field. Since there are several possible physical processes which might adequately describe this phenomenon, we choose at present to use the observational term "cancellation" and to defer introducing interpretations of the cancellation process until the observational properties are more completely established. Only a few examples of cancellation have been previously mentioned in the literature (Martin 1984; Komle 1979; Martin and Harvey 1976). In the context of this paper, the observation of cancellation is a new and significant means of differentiating pseudo ephemeral regions from real ephemeral regions.

Next we illustrate examples of pseudo ephemeral regions whose origin can be traced.

One of the most common ways for a pseudo ephemeral region to form is simply by the collision of opposite polarity fragments of network. This class of pseudo ephemeral region is expected to frequently occur in filament channels (whether or not a filament is present in the channel) because filaments and filament channels always occur at boundaries between areas of opposite polarity network magnetic fields. An example is shown in Figure 2.

Figure 2 shows the full field of the videomagnetograph in the upper section and, in the lower section, the corresponding H α filtergram. The network magnetic field is dominantly negative above the filament and positive below it. The site of the approaching network fields of opposite polarity is designated by the "1" in the lower right corner just below the filament.

Figure 3 shows a sequence of images in a limited window immediately around this pseudo ephemeral region, 1 in Figure 2. The positive pole of the pseudo ephemeral region is enclosed within the open-ended rectangle in the image at 2231. Tracing the two halves of the pseudo ephemeral regions backward in time, we see that at 1703 the pseudo ephemeral region fields originated from clumps of network field of opposite polarity. The distance between the maxima of the fragments that became the pseudo

ephemeral region was 10,000 km at the beginning of the observing day (1703). Following the sequence of images in Figure 3, forward in time, we see that the opposite polarity fields come into contact between 2124 and 2231. The relative velocity of approach of the two fragments between 1703 and 2231 is 0.5 km/sec. The fields appear to be in contact when the separation of the maxima within the opposite polarities is between 3800 and 4500 km. After their collision, it is only a short time, less than two hours and 35 minutes time gap between the last two frames, until the smaller of the two fragments of opposite polarity flux has completely disappeared.

Another common way in which a pseudo ephemeral region can form is by the encounter of one half of an ephemeral region with a fragment of network or intra-network magnetic field. An example is shown in Figure 4. A growing ephemeral region is enclosed by the oval. The opposite polarities are seen to separate from each other at a relatively rapid rate of 0.3 km/sec. The positive pole either follows or pushes an adjacent fragment of positive field into an area of strong network magnetic field of negative polarity. The negative pole of the ephemeral region during the 6 hour interval shown here is moving in the opposite direction toward a small clump of weak positive field. This weak positive fragment is either a fragment of intra-network magnetic field or a very weak fragment of network field. At 1935 we see that the negative pole of the original ephemeral region is now adjacent to both the positive fragment and a weak

negative fragment. At 2045, the negative pole of the ephemeral region and the adjacent background field of similar polarity are seen to be merging. In the final frame at 2238, the merged negative fields are abutted against the positive polarity fragment, resulting in the creation of a pseudo ephemeral region. In addition to the formation of a pseudo ephemeral region, we have illustrated in this figure the phenomenon of coalescence of magnetic fields of similar polarity, a common phenomena occurring both in the network and intra-network magnetic fields.

Figure 5 also illustrates both the coalescence of similar polarity magnetic field and the splitting of opposite polarity network field to form a pseudo ephemeral region. In this example, the videomagnetograms are shown in the form of isogauss contours. The contour levels that are shown are 10, 20, 40 and 80 gauss. Several fragments of negative polarity network or intra-network magnetic field (thinner lines) are seen in the lower part of the first frame. These fragments merge to form a single unit with simultaneous concentration of the field (second frame). Concurrently, an adjacent positive polarity fragment of network field (thicker lines) splits away from its adjoining network and in the second frame is seen to have moved toward the concentration of negative field. The result is a pseudo ephemeral region. The lower frames show a cross section of the magnetic flux (the dashed line in the upper frames) across components of the pseudo ephemeral region. In the lower part of Figure 5, the cross section profile of the positive flux in the pseudo ephemeral

region is seen to be approximately the same in both frames while the negative polarity coalescence results in an apparent increase in peak flux.

IV. ASSOCIATED H α STRUCTURES

We are also studying concurrent H α data to learn whether ephemeral regions and pseudo ephemeral regions can be distinguished from each other by means of the appearance of H α structures and to learn the direction of the component of the magnetic field parallel to the solar surface in both ephemeral regions and pseudo ephemeral regions. We have found no invariable association with any specific H α structure but we are beginning to see some trends in the data as illustrated in Figure 2. Some, but not all, opposite polarity components of ephemeral regions are seen to be connected by distinct long fibrils or arch filaments. Examples are the small region A and B in the lower left. We have found no examples where comparably long, obvious fibrils connect the opposite polarities of cancelling magnetic fields of the pseudo ephemeral regions. The cancelling pseudo ephemeral region "1" in the lower right of Figure 2 shows a small filament or fibril dividing rather than connecting the opposite polarities. Such dividing fibrils, however, are not invariably observed. No unique structure is associated with cancelling feature "2" in the right side of this figure. To date, we have found no definite

signatures whereby ephemeral regions or pseudo ephemeral regions can always be identified in H α .

Some pseudo ephemeral regions are associated with microflares. An example is shown in Figure 6, a superposition of H α and 10 gauss contours. A new, small ephemeral region is labelled ER in the first frame. As it grows, the negative (dashed lines) of the ephemeral region collides with adjacent positive polarity (solid lines) network and a small two point micro-flare occurs at the junction of the ephemeral region and opposite polarity network. This relationship to the associated magnetic fields is exactly the same as described by Marsh (1978) in his study of ephemeral region flares. We only emphasize the newly recognized process of cancellation. We also call attention to the fact that the time scale of cancellation is on the order of a few hours for small-scale quiet sun features while microflares are typically seen only for a few minutes in the chromosphere. However, if microflares are analogous to larger flares, the coronal part of the microflares could last ten times longer than the chromospheric part of the event.

V. ASSOCIATED FEATURES AT OTHER WAVELENGTHS

We anticipate that pseudo ephemeral regions, during the stage of cancellation, may have more or less steady signatures at radio, ultra-violet and soft X-ray wavelengths, if the cancellation represents slow reconnection or any other type of slow conversion of magnetic energy to other forms of energy. In addition, at these wavelengths, we expect transient signatures associated with microflares.

The cancelling pseudo ephemeral regions are excellent new candidates for a direct association with many X-ray bright points which are also known to produce tiny flares (Golub et al. 1974). If our hypothesis is correct, many more X-ray bright points may be associated with pseudo ephemeral regions than with real ephemeral regions. The association with ephemeral regions would in many cases still be very close since many pseudo ephemeral regions are the consequence of one-half of an ephemeral region colliding with network magnetic fields. Additionally, we do not rule out the possible association of some X-ray bright points with real ephemeral regions; we only note some new possible associations. The circumstance of two mutually cancelling ephemeral regions would be an excellent candidate for association with X-ray bright points, although it is relatively rare.

A principal difficulty with previous hypotheses (Harvey et al. 1975; Golub et al. 1977; Golub 1980) of a one to one association between ephemeral regions and X-ray bright points was the uncertainty of whether there were sufficient small-scale, uncounted, ephemeral regions to match one for one with the X-ray bright points. The existence of cancelling pseudo ephemeral regions, as well as many cancelling magnetic features which would not be mistaken for ephemeral regions, offers a sufficient number of distinct magnetic features to account for all X-ray bright points.

Our hypothesis that many X-ray bright points may not be the counterparts of ephemeral regions is in direct contradiction to the assumption of Golub (1980) that X-ray bright points represent emerging magnetic flux. We emphasize the importance of the simultaneous observation of quiet sun magnetic fields at X-ray, optical and radio wavelengths.

VI. DISCUSSION

Out of hundreds of ephemeral regions observed from birth, we have not yet observed a single one in which the growing and separating opposite polarity fields reverse their direction of motion, come together, and disappear, except in one case in which there is also evidence that two ephemeral regions mutually cancel each other. Thus we must suspect that

the cancellation of an ephemeral region within itself is something that either does not happen or is extremely rare. We therefore have a sharp distinction between ephemeral regions and pseudo ephemeral regions that can be used to distinguish these two phenomena when the origin of the fields has not been observed. A few hours observation is sufficient to determine if the fields are moving together and cancelling, are stationary or unchanging, or growing and separating. However, we also need to point out that some circumstances are ambiguous, especially when an ephemeral region grows in the middle of network magnetic fields. In these cases, we usually observe a more rapid growth in the half of the ephemeral region that is the same polarity as the network and less rapid apparent growth in the half of the ephemeral region that is opposite in polarity to the network. We interpret the slowly growing half as due to growth and concurrent cancellation of that half of the ephemeral region. In such cases it is clear that the growth rate of the ephemeral regions must exceed the rate of flux loss due to cancellation. Otherwise ephemeral regions would not be found amidst the network fields. Outside of the network there is no requirement for the ephemeral region growth rate to exceed the flux loss of cancelling features in general. The consequence of these differences is that we expect to find a minimum size for ephemeral regions that occur coincident with network. Additionally, the lifetime of such ephemeral regions occurring in or very near network are thus greatly shortened in comparison to ephemeral regions that do not

occur either in or very close to network fields.

Pseudo ephemeral regions cover the entire range of total fluxes found in ephemeral regions. This is not surprising since elements of opposite polarities covering a wide range of magnetic flux may move together on the sun. In contrast to the ephemeral regions, pseudo ephemeral regions do not originate as new bipolar fields. They constitute opposite polarity fragments of magnetic fields that have previously constituted network and intra-network magnetic fields, separate halves of ephemeral regions or any combination of these sources. They are brought together by the motions of the network magnetic fields along the boundaries of supergranule cells, by convection within the supergranule cells, by the growth and separation of the opposite polarity halves of real ephemeral regions. Even though the opposite polarity halves of the pseudo ephemeral regions originate from separate sources, it is noted that they often have some properties in common:

- (1) cancellation of the opposite polarity fields when they come within a critical distance from each other (cancellation is here defined as a mutual decrease of the magnetic field of both features)
- (2) increasing gradient between the opposite polarities
- (3) continued motion of the opposite polarities towards each other as the cancellation proceeds, resulting in a continued concentration

and reduction of the total flux of the pseudo ephemeral regions.

The above properties of pseudo ephemeral regions are all characteristics not shared by isolated ephemeral regions. Since they are evolutionary characteristics, one or more hours of observation may be required to correctly identify whether any apparent bipole is a new ephemeral region or whether it is a pseudo ephemeral region. However, the gradient of the magnetic field in the region between opposite polarities can be used to identify some cancelling pseudo ephemeral regions with only a single magnetogram. In Figure 7, we illustrate both a real ephemeral region, in the middle of the field, and a pseudo ephemeral region, in the lower right. In the second frame at 0027 UT, the steepness of gradient of the pseudo ephemeral region alone is sufficient evidence that this feature is not a real ephemeral region. In Big Bear magnetograms the gradient measured in the polarity inversion line between points separated by about 2000 km will very seldom exceed 0.01 gauss/km in a real ephemeral region.

We also note that the maxima in each polarity of the magnetic field in the ephemeral region are centered nearly symmetrically within the contours of lower field strength. In the pseudo ephemeral region, we observe that the maxima in opposite polarities crowd towards each other as the gradient of the field increases. Because of these very obvious characteristics, it is easier to identify pseudo ephemeral regions than real ephemeral regions, especially in data sets with low time resolution.

VII. DIRECTIONS OF CONTINUING RESEARCH

The existence of the pseudo ephemeral regions presents additional factors not accounted for in previous estimates of the number of ephemeral regions that truly exist on the whole sun at any one time. We now recognize that the number of ephemeral regions, that can be counted on a set of data, depends on:

- (1) the spatial resolution of the telescope used,
- (2) the instrumental sensitivity coupled with exposure time,
- (3) the spacing of observations (continuous time-lapse observation, one per minute, affords the ability to recognize more ephemeral regions than more widely spaced observations such as once an hour or once a day),

and the following solar factors:

- (4) the distribution of network magnetic fields over the solar surface,
- (5) the time in the solar cycle.

The primary limitation in making new estimates is item (4) above, our yet incomplete knowledge of the interaction of ephemeral regions with the network. We need to know whether or not there is any preference for ephemeral regions to occur in the presence of network magnetic fields as

we might suspect from the results of Garcia de la Rosa (1983), Martin et al. (1983), Gaizauskas et al. (1983), and Liggett and Zirin (1984). We also need to know the distribution of the rates of separation of the opposite polarity components of ephemeral regions and the rates of cancellation of ephemeral regions. All of these unknowns limit our ability to extrapolate to find the population of ephemeral regions on the full sun from limited field observations. However, these are only temporary limitations which we are attempting to solve through the continued analyses of the data on hand and the acquisition of additional data.

VIII. CONCLUSIONS

We have found that ephemeral active regions can be differentiated during their growth stage from other apparent bipoles on the sun by their characteristic evolution. Ephemeral regions are new bipolar magnetic fields that originate with their opposite polarities adjacent or very close to each other. Ephemeral regions always show a growth stage in which the maxima of the opposite polarity fields separate as a function of time.

Pseudo ephemeral regions can either be unchanging, stationary bipolar fields or opposite polarity fragments of magnetic field which have moved together and exhibit cancellation of both polarities at their mutual boundary. Our study of limited fields of the quiet sun shows that pseudo

ephemeral regions with cancellation outnumber pseudo ephemeral regions with little or no cancellation, and pseudo ephemeral regions outnumber the real ephemeral regions by about a factor of two.

The origin of the magnetic fields of pseudo ephemeral regions can be from:

- (1) fragmentations of very old network magnetic fields,
- (2) from relatively new concentrations of intra-network fields,
- (3) from halves of ephemeral regions which are no longer recognizable,
or
- (4) from any combination of the previous three possibilities.

In H α the opposite polarity fields of some ephemeral regions are connected by fibrils and the opposite polarities of some pseudo ephemeral regions are divided by small filaments. Most ephemeral regions and pseudo ephemeral regions are difficult to recognize in H α photographs without the aid of magnetograms. Micro flares occur in at least some pseudo ephemeral regions. This evidence of energy release is consistent with our hypothesized association of many X-ray bright points with pseudo ephemeral regions and other cancelling magnetic features. We further expect distinct radio signatures from pseudo ephemeral regions, at least during the occurrence of microflares.

Acknowledgments

J. Wang and Z. Shi express their appreciation to Dr. H. Zirin for the extension of their visit at Caltech. The authors thank A. Patterson for many helpful discussions on observing technique and for presenting the paper at MSFC. The contribution of S. F. Martin to this paper was supported by NASA Contract NAGW 255. Facility and observational support at Big Bear Solar Observatory was provided under NASA Contract NGL 05 002 034 and NSF Grant ATM 8211002.

References

- Gaizauskas, V., Harvey, K. L., Harvey, J. W., and Zwann, C., 1983, Ap. J. 265, 1056.
- Garcia de La Rosa, J. I., 1983, Solar Phys. 89, 51.
- Golub, L., Krieger, A. S., Silk, J. K., Timothy, A. F., and Vaiana, G. S. 1974, Ap. J. 189, L93.
- Golub, L., Krieger, A. S., Harvey, J. W., and Vaiana, G. S. 1977, Solar Phys. 53, 111.
- Golub, L. 1980, Phil. Trans. R. Soc. Lond. A297, 595.
- Harvey, K. L., and Martin, S F. 1973 Solar Phys. 32, 389.
- Harvey, J. W., Harvey, K.L., and Martin, S. F. 1975, Solar Phys. 40, 87.
- Komle, N. 1979, Solar Phys. 64, 213.
- Liggett, M., and Zirin, H. 1984, Solar Phys., submitted.
- Marsh, K. 1978, Solar Phys. 64, 93.
- Martin, S. F. 1984, Proceedings of the Symposium on "The Small-scale Dynamics in Stellar Atmospheres," Sacramento Peak Observatory, 25-29 July 1983.
- Martin, S. F., Dezso, L., Antalova, A., Kucera, A., and Harvey, K. L., 1983, Adv. Space Res. 2, 39.

Figure Captions

Fig. 1 - The evolution of the new ephemeral active region enclosed in the oval (1930) can be compared to the pseudo ephemeral region enclosed in the rectangle in the upper right. As the ephemeral region grows, the pseudo ephemeral region exhibits "cancellation," the mutual loss of magnetic flux in closely spaced features of opposite polarity. Another example of cancellation of a small positive polarity fragment (white) of field with negative polarity (black) network magnetic field is enclosed within the corner of the rectangle to the lower right of the ephemeral region.

Fig. 2 - Pseudo ephemeral region 1 shows a small filament or fibril dividing the opposite polarity fields as cancellation occurs. Cancelling feature 2 shows no unique signature in H α . Non-cancelling bipolar regions A and B reveal fibrils or arch filaments connecting their opposite polarities.

Fig. 3 - The open-ended rectangle marks the site of a pseudo ephemeral region which originates from opposite polarity fragments of network coming together. The H α structures corresponding to this pseudo ephemeral region and other features are shown in Figure 2.

Fig. 4 - The negative half (black) of a growing, separating ephemeral region (in oval) is seen moving toward opposite polarity (white) network magnetic field and forms the "pseudo ephemeral region" within the open ended rectangle at 2238. The positive half of the ephemeral region pushes an adjacent positive field fragment toward a large clump of negative polarity network where field "cancellation" has the opportunity to begin.

Fig. 5 - A pseudo ephemeral region is formed from the coalescence of positive polarity magnetic field fragments and the simultaneous approach of a network fragment of opposite polarity. Contour levels are 10, 20, 40, and 80 gauss.

Fig. 6 - The positive (dashed contour line) and negative (solid contour line) poles of the small ephemeral region (ER) separate from each other between 1836 and 2135. After the positive pole collides with a fragment of negative polarity network, a two point microflare occurs (2135). Meanwhile, the negative pole of the ephemeral region undramatically merges with network of the same polarity.

Fig. 7 - An important distinguishing characteristic of pseudo ephemeral regions is increasing field gradient. The appearance of the pseudo ephemeral region is shown in contrast to the real ephemeral region which appears in the center of the second frame. The magnetic field near

the polarity inversion line separating the two halves of the pseudo ephemeral region is characterized by higher gradient than near the polarity inversion line in the real ephemeral region. Contour levels are 5, 10, 20, 40, and 80 gauss.

S. F. MARTIN: Solar Astronomy 264-33, California Institute of Technology,
Pasadena, California 91125

S. H. B. LIVI: Instituto de Fisica, Universidade Federal do Rio Grande do Sul,
90,000 Porto Alegre, Brazil

J. WANG: Beijing Observatory, The Academy of Sciences, Beijing, China

Z. SHI: Beijing Observatory, The Academy of Sciences, Beijing, China

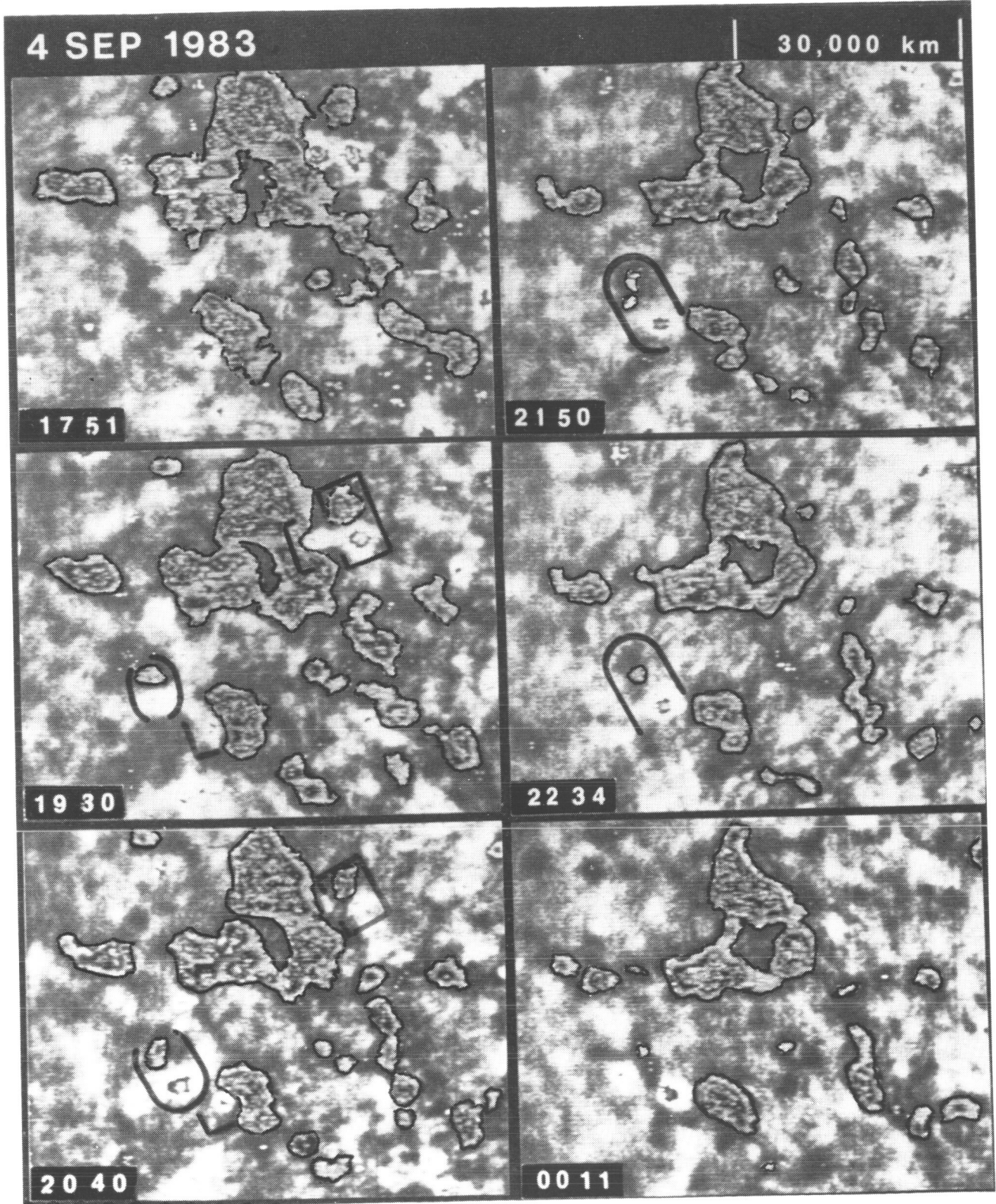


Figure 1

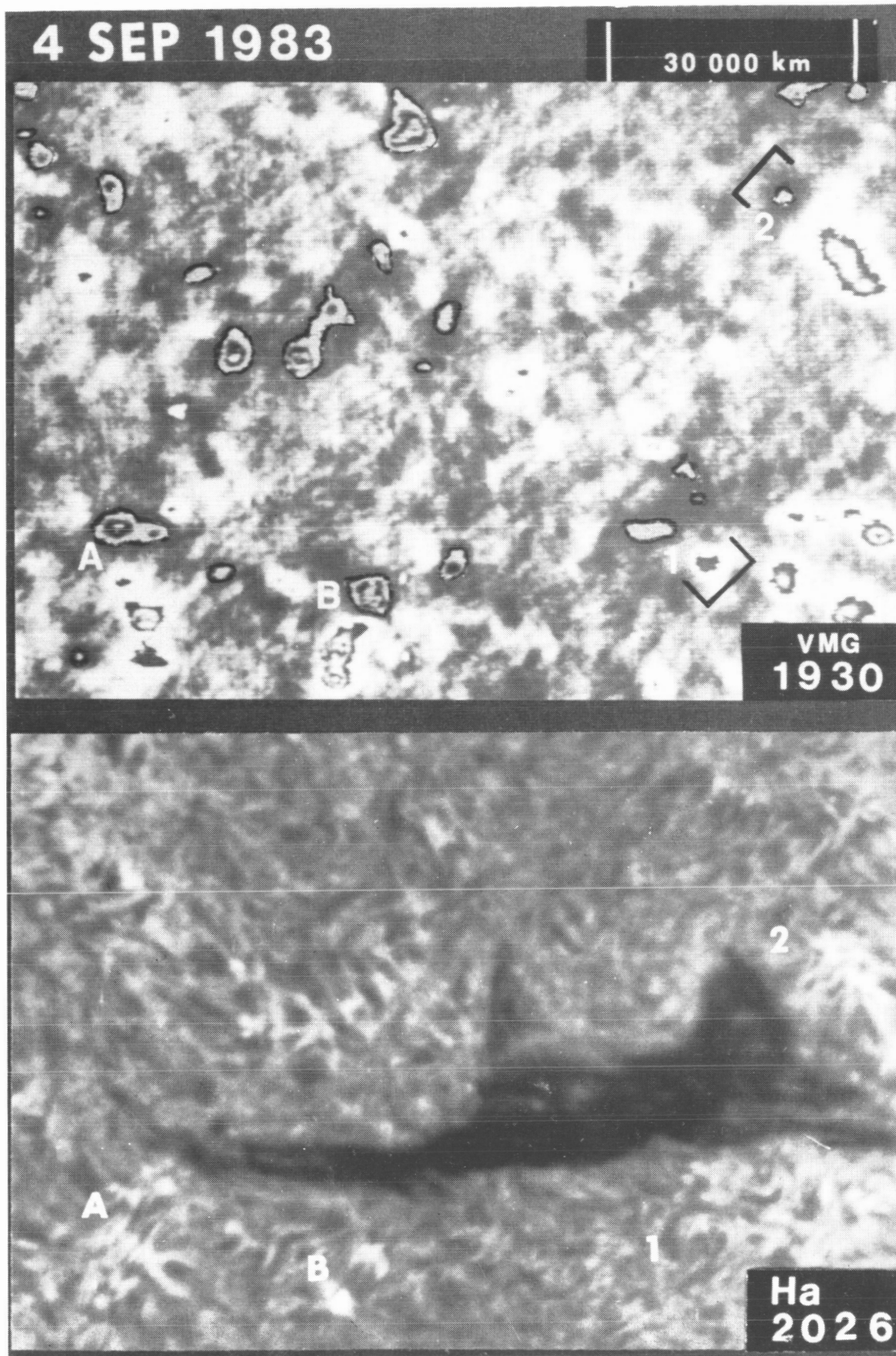


Figure 2

4 SEP 1983

30 000 km

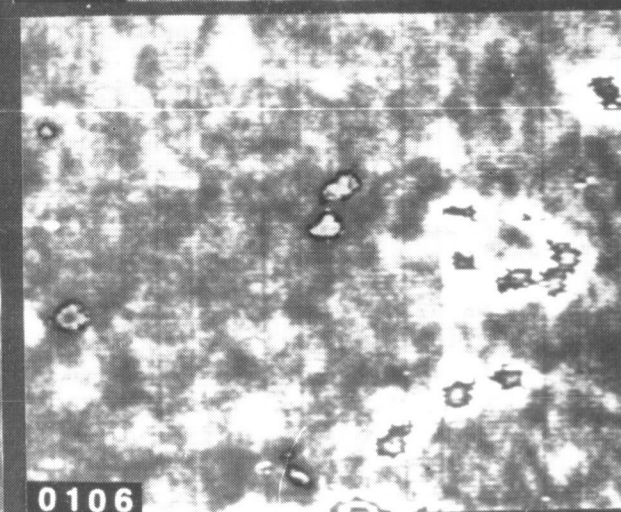
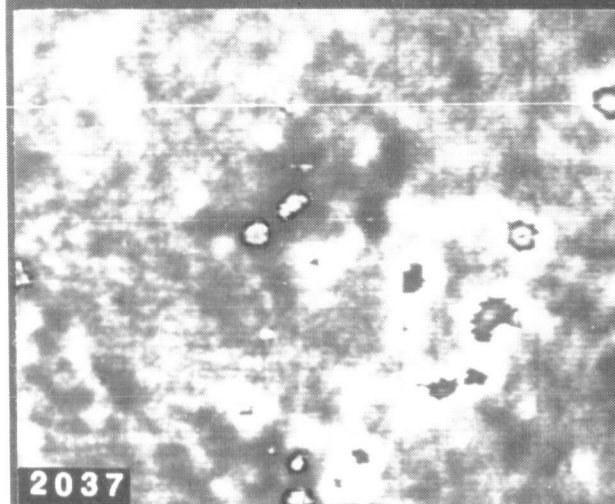
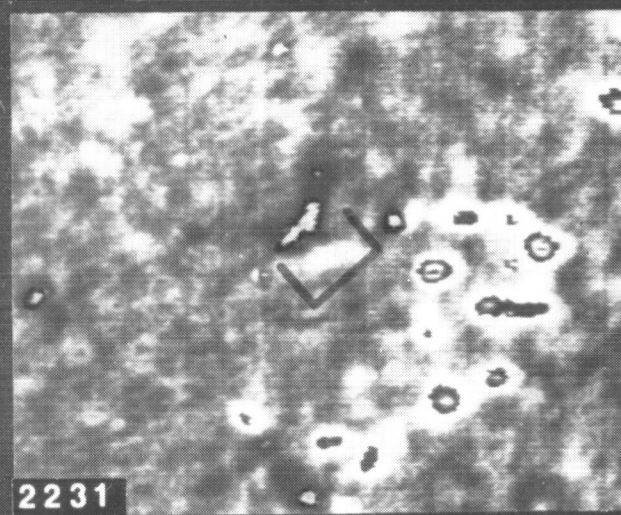
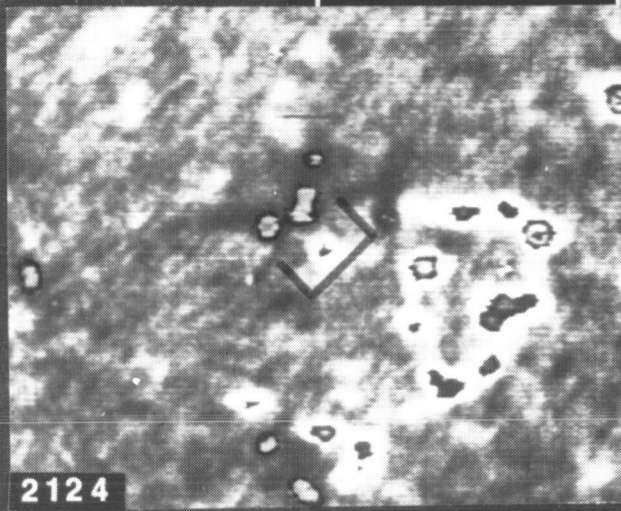
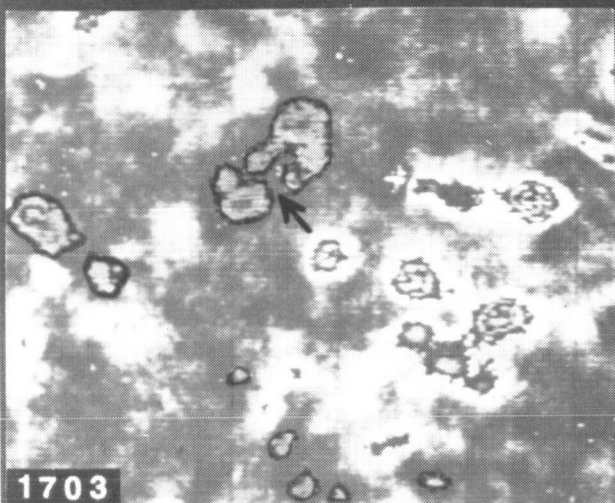


Figure 3

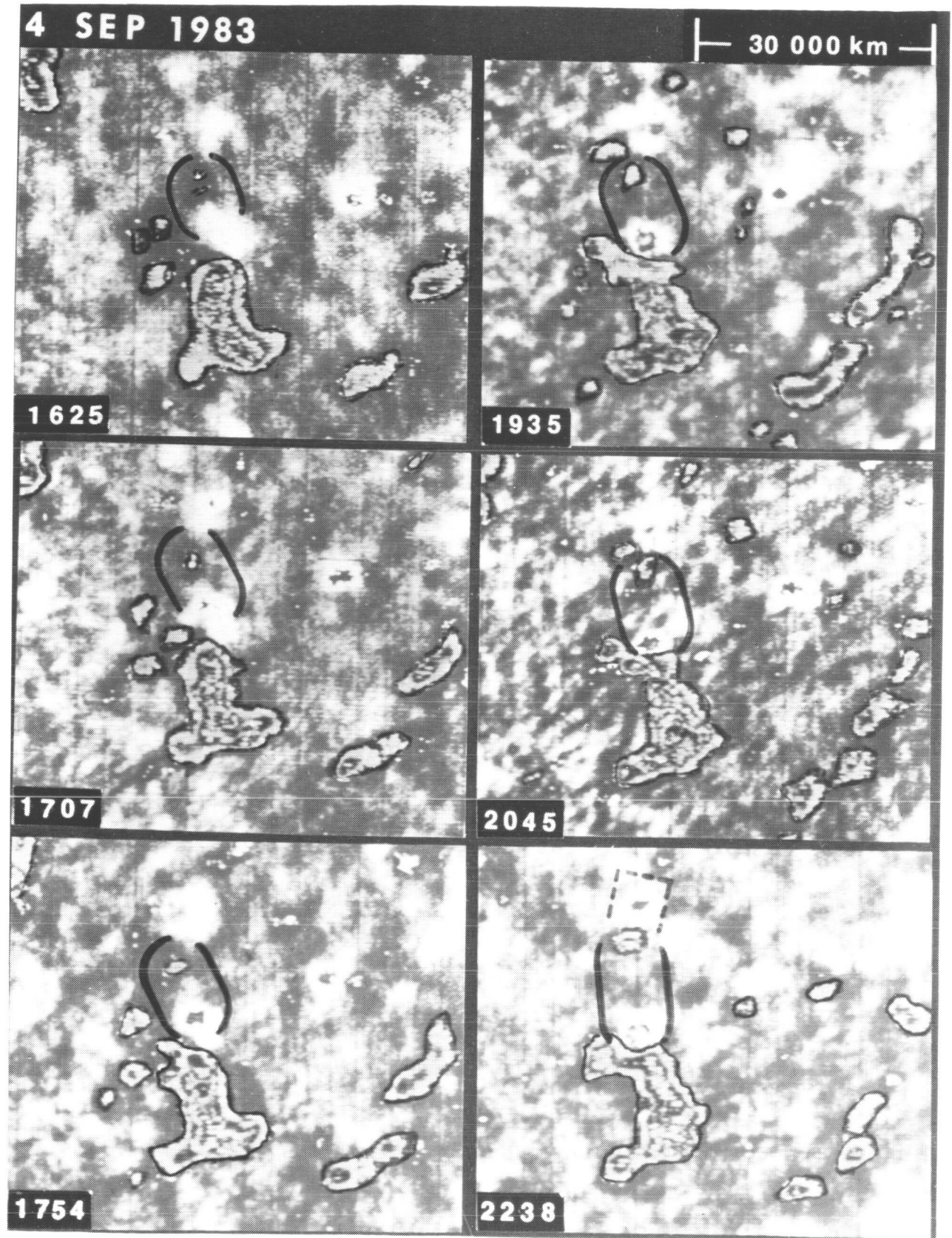
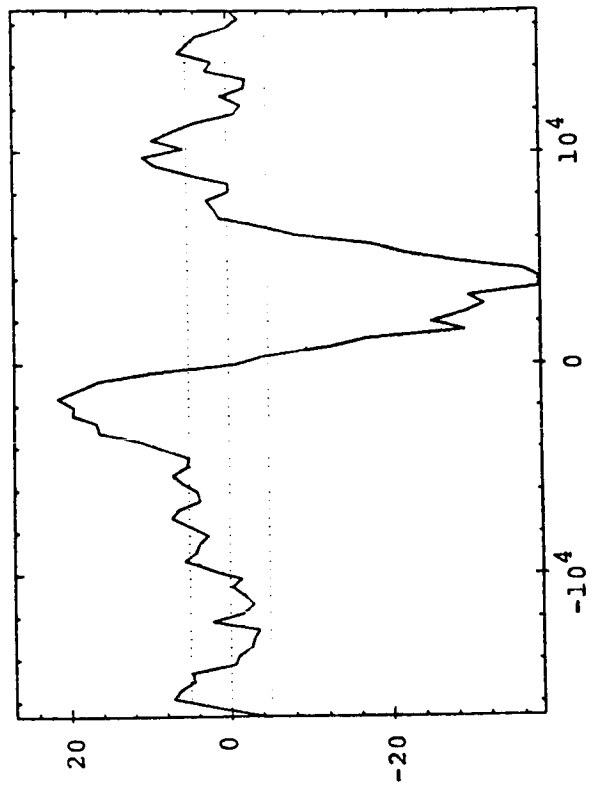
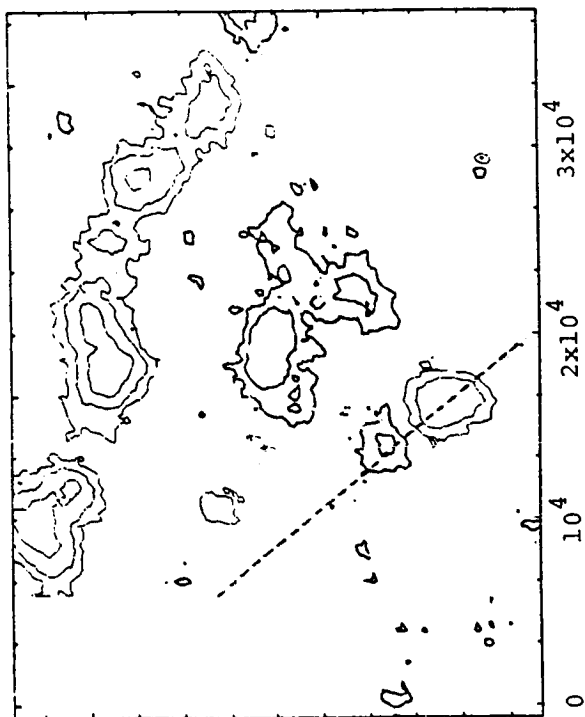


Figure 4

2135 Oct. 14



1836 Oct. 14

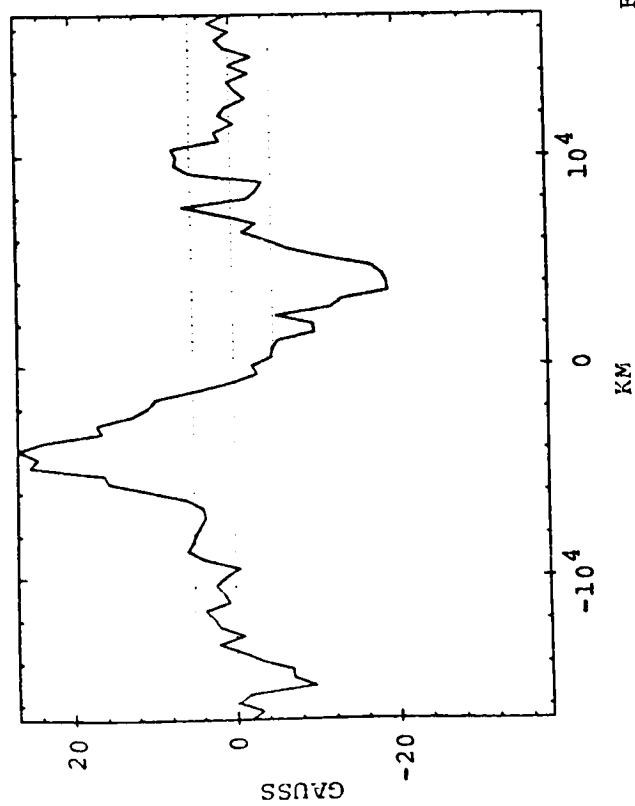
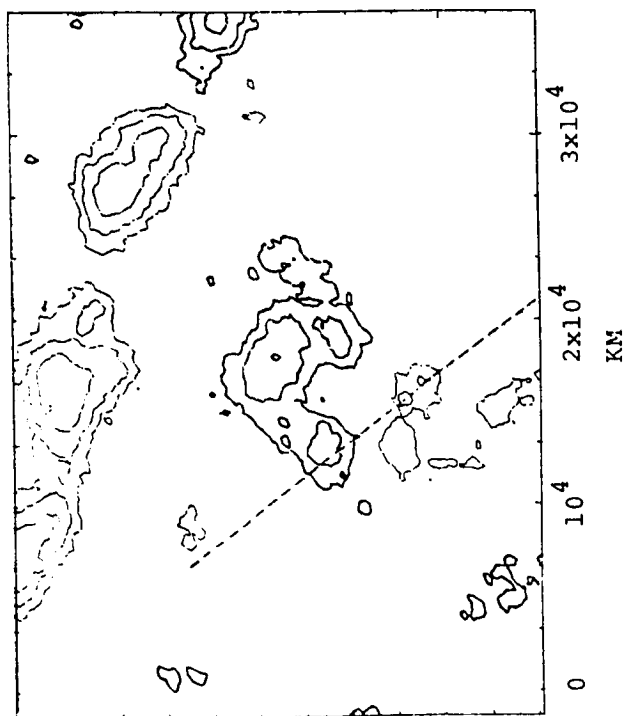


Figure 5

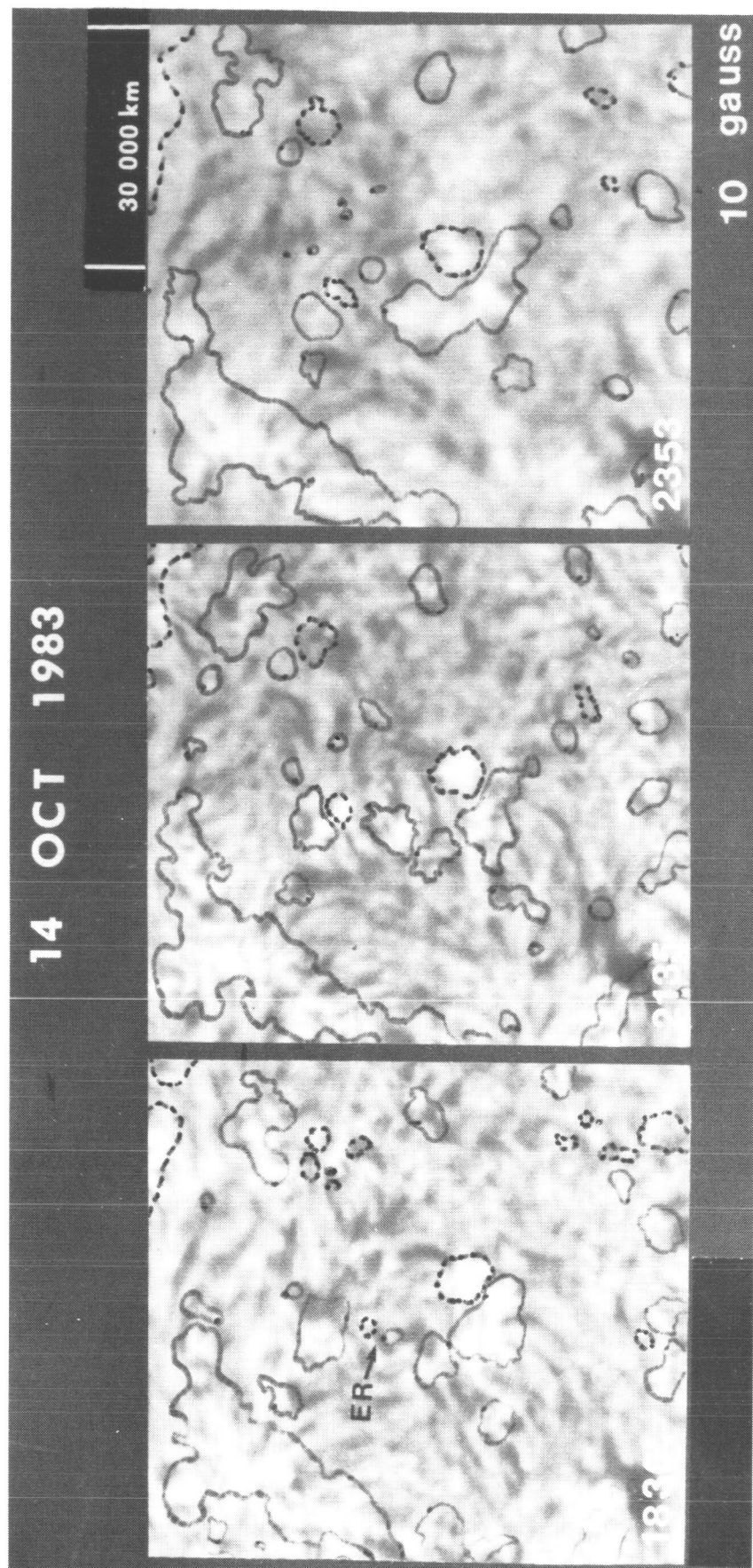
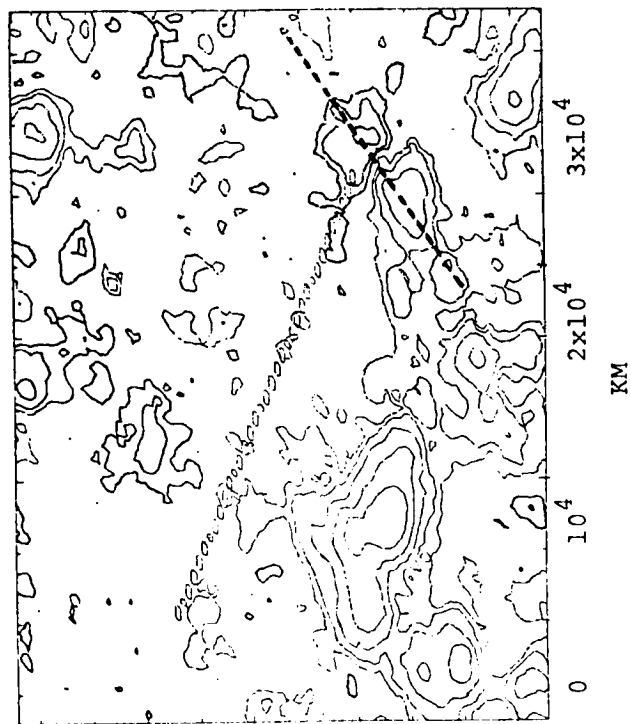


Figure 6

2251 Sep. 4



0027 Sep. 5

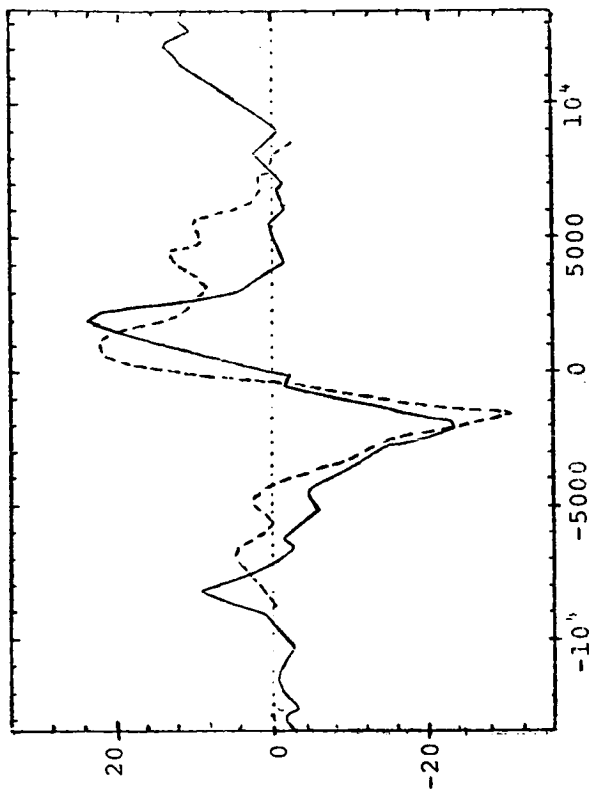
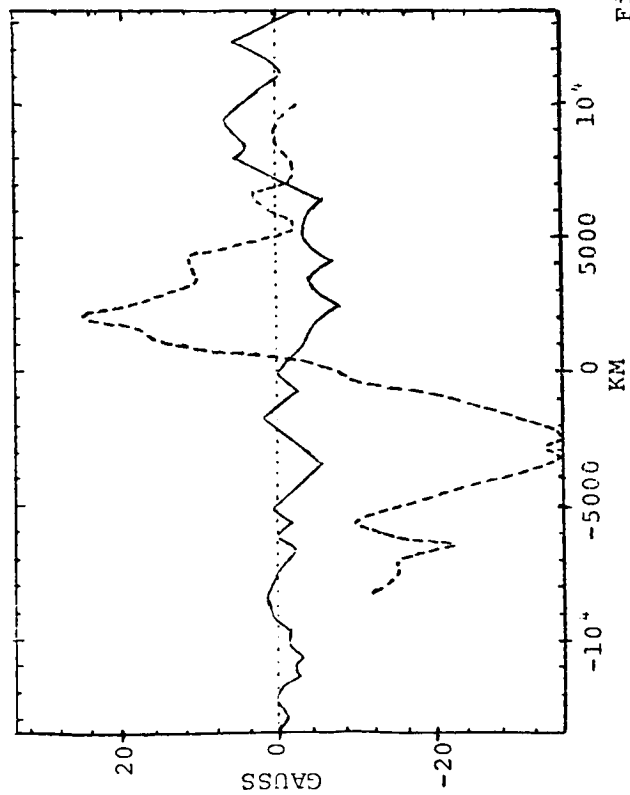
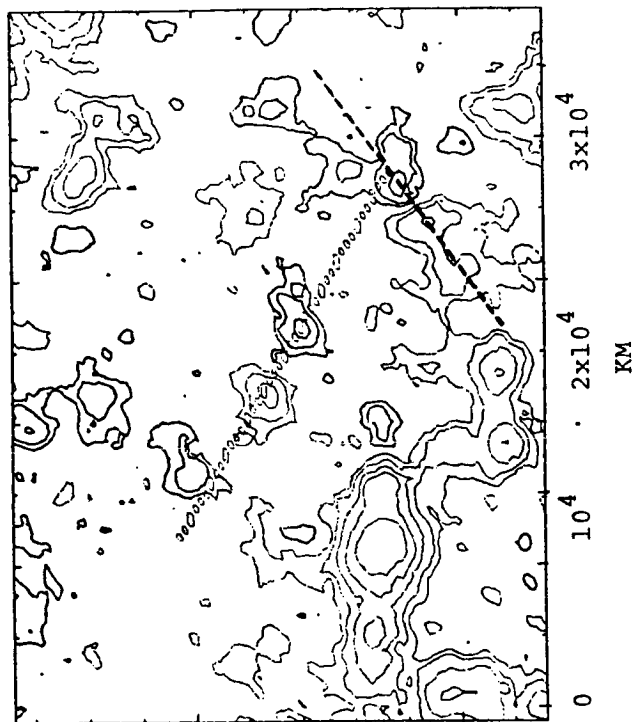


Figure 7

Evidence for Submergence of Magnetic Flux
in a Growing Active Region

D. M. Rabin, NRC/Marshall Space Flight Center
R. L. Moore, NASA/Marshall Space Flight Center
M. J. Hagyard, NASA/Marshall Space Flight Center

In NOAA Active Region 2372 (April 1980), 4×10^{20} maxwell of magnetic flux concentrated within a 30" circular area disappeared overnight. Vector magnetograms show that all components of the magnetic field weakened together. If the field had weakened through diffusion or fluid flow, 80% of the original flux would still have been detected by the magnetograph within a suitably enlarged area. In fact there was at least a threefold decrease in detected flux. Evidently, magnetic field was removed from the photosphere.

Since the disappearing flux was located in a region of low magnetic shear and low activity in $H\alpha$ and $L\alpha$, it is unlikely that the field dissipated through reconnection. We argue that the most likely possibility is that flux submerged. Our observations suggest that even in the growth phase of active regions, submergence is a strong process comparable in magnitude to emergence.

This research is supported by the National Research Council and by the NASA Office of Solar and Heliospheric Physics.

SECTION 4

TECHNIQUES FOR DATA DISPLAY

SOLAR PHYSICS APPLICATIONS OF COMPUTER GRAPHICS AND IMAGE PROCESSING

Martin D. Altschuler
Department of Radiation Therapy
University of Pennsylvania School of Medicine
3400 Spruce Street
Philadelphia, Pennsylvania 19104

ABSTRACT:

Computer graphics devices coupled with computers and carefully developed software provide new opportunities for the research scientist to achieve insight into the geometry and time evolution of scalar, vector, and tensor fields and to extract more information more quickly and cheaply from the same image data.

Two or more different fields (such as, temperature, density, and velocity) which overlay in space can be calculated from the data (and the physics), then displayed from any perspective, and compared visually. The maximum regions of one field can be compared with say the gradients of another. Time-changing fields can also be compared.

Images can be added, subtracted, transformed, noise filtered, frequency filtered, contrast enhanced, color coded, enlarged, compressed, parametrized, and histogrammed, in whole or section by section. Solar physicists forever show flare movies of several different spectral lines together with white light images. Today it is possible to process multiple digital images to reveal spatial and temporal correlations and cross-correlations. Moreover, data from different observatories taken at different times can be processed, interpolated, and transformed to a common coordinate system.

Statistics can be performed pixel by pixel or segment by segment, with histograms displayed on the screen in real time if desired. This statistical information can be used for adaptive filtering, to study a field variable over different image regions of varying contrast.

At present the development of a facility for research-quality graphics and image processing is a considerable investment in time and money. Graphics software is still very machine dependent and not very portable (although recently giant steps toward adopting GKS standards have been taken). Although many graphics and image-processing systems come with a subroutine library, the subroutines are rarely sufficient for specific research problems, so that a considerable effort is needed to develop high level software. Solar physicists tend to be rugged individualists who work alone or in small groups; few have the time or expertise to do their own research and also develop sophisticated software for image processing and graphics.

I recommend that several observatories standardize on a common minicomputer and graphics system, that a separate system be purchased for each observatory, and that an intermediate level of

image processing and graphics software be developed and maintained jointly by all the participating groups. From this shared software library, each researcher can more easily develop the particular application software for his own research problem.

1. Introduction

We are entering a new age of pictorial communication in which photographic images, realistic simulations, sketches, contour plots, engineering drawings, stereo projections, cinema, and so on, will be created, displayed, processed, edited, and transmitted with as much ease as printed text. In this paper we are concerned with the use of this new medium as a tool in scientific research. Solar physics in particular is probably one of the few branches of science where image processing and computer graphics can potentially offer tremendous advantages in (1) the collection, archiving, and retrieval of vast amounts of image data, (2) the visual comparison of many different types of data-derived information, including complex velocity and magnetic vector fields, (3) the transformation of data collected over different times, thus different perspectives on the rotating sun, to the same reference frame for viewing, (4) the creation of synoptic data bases which can be readily transferred from place to place so that different research groups can test their theories, search for correlations, and put their high resolution observations into global context.

I will partition this paper into three parts, of unequal size and depth of detail, addressing the questions:

- (1) What is computer graphics and computer imaging?
- (2) What can it do for solar physics?
- (3) How should graphics and image processing facilities be set up in the field of solar physics?

2. Overview of Computer Graphics and Image Processing

There are a number of excellent introductory texts that can painlessly introduce the working solar physicist to the ideas of computer graphics and image processing. I would recommend at least the following:

Principles of Interactive Computer Graphics, 2^d edition,
W.M. Newman and R.F. Sproull, McGraw-Hill, New York, NY, 1979.

Fundamentals of Interactive Computer Graphics, J.D. Foley and A. Van Dam, Addison-Wesley, Reading, MA, 1982.

Algorithms for Graphics and Image Processing, T. Pavlidis,
Computer Science Press, Rockville, MD, 1982.

Computer Image Processing and Recognition, E.L. Hall, Academic Press, New York, NY, 1979.

Image Reconstruction from Projections, G.T. Herman, ed., Topics in Applied Physics, Vol. 32, Springer-Verlag, Berlin, 1979.

We can begin our thinking about graphics and image processing by considering mappings from one set of data to another.

<u>input</u>	<u>output</u>	<u>process or operation</u>	<u>application</u>
number	number	scalar function	analog devices
function	function	transform	signal processing
image	image	2-D transform	image processing
model	image	projection, illumination, texture	simulation
image	model, class	pattern searching	pattern recognition

Image processing also includes the reconstruction of a 2-D cross-section from its one-dimensional projections, and the reconstruction of a 3-D object from its 2-D cross-sections (or directly from its 2-D projections), with the result displayed as a solid in perspective. These applications, important in x-ray computed tomography and magnetic resonance imaging, are often referred to as image reconstruction methods.

Another way to think about graphics and image processing is in terms of the types of images and the transformations between them. The four types of images are: (1) full gray-scale images, (2) binary (or black-and-white) images, (3) continuous curves or contours, and (4) piecewise-continuous curves, such as polygons or sets of points (cf. Pavlidis, 1982, cited above).

The first two types of images are raster images because the image is made up of a raster of small rectangular area elements (pixels); the intensity (gray level) of every pixel is stored as a number in an image memory. The last two types of images are vector images; the sequence of vectors is stored in memory as a sequence of point coordinates.

To display a raster image, every word in an image memory must be interrogated. The size of the image memory word depends on the number of gray levels (intensity shades) that each pixel can display. For example, to display 256 shades of gray, we need a word of 8 bits (called a byte), so that a word in image memory with value 00000000 sets its corresponding pixel to black, and a word of value 11111111 sets its pixel to white (the 255th gray level, in accord with the binary number system). To display a black-and-white image, however, we need only one bit in a word. Thus an image memory of 512 x 512 bytes can display an image of 512 x 512 pixels with 256 gray levels, or else 8 black-and-white images of 512 x 512 pixels. Black-and-white raster images of high spatial resolution (small pixel size) can also be used to store vector information. When a raster device is capable of displaying vector graphics, the machine is sometimes called a raster graphics display. For much solar physics work, multi-gray-level images are the raw data, so that raster displays with raster graphics capability are recommended.

Color displays can be generated from three separate gray-level or black-and-white raster images. Each color is generated by its own cathode beam controlled by its own image memory. With 256 gray levels in each color, realistic color images (essentially color photography) can be achieved. With fewer gray levels, we can still achieve color maps and overlays, and cartoon-quality images. Since color images are simply 'gray-level images with triple memory', we need not discuss color images separately. In solar physics research, color is useful to distinguish regions of different magnetic polarity, or regions of upward versus downward fluid flow, or to color-code regions of different emission, etc.

A transformation from one multi-gray-level image to another falls in the category of image processing. Image processing techniques are used for a number of purposes, such as

- (1) Contrast enhancement: Suppose your images are over or under exposed so that the region of interest (say a sunspot umbra) extends over only a relatively small range of gray levels. This technique extends the contrast range for the region of interest. In the extreme case, all intensities exceeding the maximum in the region of interest are set to white and all those below the minimum are set to black, with the region of interest scaled to cover the remaining 254 gray levels.
- (2) High-frequency spatial filtering: Sometimes an image is very noisy, or interference fringes or pixel discretizations are distracting. Then all high spatial frequencies or specific spatial frequencies can be filtered.
- (3) Low-frequency spatial filtering: If you are interested in rates of change of intensity in an image, for example when tracing chromospheric filaments or the boundaries between active regions, or outlining the gradients on a magnetograph display, then you can filter out the low spatial frequencies and enhance the appearance of boundaries and sharp gradients.
- (4) Image differencing: If you want to detect changes in intensity or field from one day to the next, you can difference two images and shade in or enhance the regions of change.
- (5) Image averaging: If you take images every minute or so but are worried about atmospheric distortions, you can average the images together to improve the statistics and decrease noise.

Transformations between black-and-white images are of use in following the changes in shape, area, or position of certain specified image features. It might be of interest, for example, to follow the shape of a sunspot umbra over time. We first transform coordinates so as to view the umbral region from the same perspective, and then display the umbral area observed at each time in a different color overlay. Alternatively, only the changes in the umbra over time need be colored.

Transformations between continuous-curve images are used in map contouring and in the study of changes in intensity as indicated by contour outlines. The study of flare activity in a region over several minutes can be studied in this way.

Piecewise-continuous curves, such as polygons, are used for rough outlining of contours when computer memory is limited, and for editing computer graphic images (interactive graphics).

Multi-gray-level images are transformed to black and white images when a specific segment or image area is to be detected or emphasized. For example, we can examine a set of multi-gray-level images in hydrogen-alpha, set all pixels brighter than a certain threshold intensity to white (on) and all other pixels to black (off). A black and white image will then show just the flare regions.

When we have black and white images, we can transform to a continuous or piecewise-continuous vector image by storing just the points on the contour or just the skeleton of the region (cf. Pavlidis, 1982, cited above).

In general, when we transform from multi-gray-level images to black and white images to continuous-curve images to piecewise-continuous images (points and polygons), we are usually involved in processes of pattern recognition. We are usually recognizing, detecting, or emphasizing certain critical features of an image, and are unconcerned about the background. Since it requires considerable computer memory to store multi-gray-level images, the more pattern recognition that can be done, the smaller the required memory, the smaller the data base, and the less the access time to retrieve the data.

Piecewise-continuous contours may be drawn as overlays on a multi-gray-level image by a user with a lightpen or trackball. We might want to smooth a contour, or have the machine search within a hand-drawn contour until it finds a specified intensity level in the image and then sketch a smooth isocontour. The output desired is a continuous-curve image. A transformation from polygon contours to continuous contours is called interpolation.

The transformation from a vector contour image to a black and white raster image is called contour filling or shading. A shaded image sometimes give a better feel for the data than the contours alone.

The transformation from a black and white image to a multi-gray-level image may be done by (1) dithering, in which spatial resolution is traded for gray level information, (2) the introduction of false color, and (3) the use of artificial illumination and shading. In radiological analysis, a series of CT or MR scans is first used to derive the cross-sectional reconstructed images. The bones are thresholded in each cross-section to form a series of black and white images, and a computer then creates an artificially illuminated and shaded 3-D perspective multi-gray-level image of the bone surface as it might appear in 3-D. The object can be rotated at the computer console to change the perspective. In solar physics work,

different synoptic data bases might be interrogated and a synthesized multi-color image created to provide a multi-data perspective of a solar feature.

In general, when we transform from polygon-contour images to continuous-contour images to black-and-white images to multi-gray-level images, we are involved in computer graphics processes. Computer graphics allows us to visualize a data base in context or view a model with realism. For solar physics work, the primary application should be to allow the user to rapidly understand complex data bases pictorially, without wading through mountains of number tables or examining hundreds or raw photographic images.

3. Computer Graphics as a Tool for Solar Physicists

In general there are five basic activities that a solar physicist can expect to accomplish with computer graphics and image processing capability: (1) processing raw images, (2) comparing different data or information, (3) transforming data or information to different coordinate systems, (4) creating data bases from processed or unprocessed data, and (5) interrogating and visualizing the information in data bases.

3.1 Processing raw images.

A large collection of solar images, such as hydrogen-alpha pictures, needs to be stored (archived), indexed for searching, retrieved upon request, and transmitted to various devices or to other users. At present too much memory is required to store many images comfortably in digital form. (With optical read-only memories, however, this situation may soon change.) Images can be stored in analog form on video disks or video tapes. To achieve sampling to better than 512 pixel resolution, parts of a solar image can be stored in separate video frames. When a video frame is requested from an index, the frame is retrieved and digitized by a fast A/D (analog to digital) converter. In some image processing systems, a separate memory plane is dedicated to each 512 x 512 pixel image, so that one can store in digital form as many 256 gray-level images as there are memory planes. This allows rapid processing of several images, for example, image differencing or averaging. By retrieving images from video tape, digitizing the images, and applying image-processing, thresholding, or pattern-recognition algorithms, the user can rapidly study selected solar features over a period of time. Pattern recognition usually requires more than simple image processing, so that rather complex, and usually heuristic, algorithms must be written for each specific application.

A collection of solar pictures can also be studied interactively. The user can outline boundaries on the solar disk, either manually by light pen or by calling edge-detecting

algorithms. Statistics related to brightness or solar area of selected features (say flares or sunspots) can be shown on the screen. If the central meridian longitude and inclination of the sun are known, and the solar disk and the rotation axis outlined, then the latitude and longitude of any point on the solar image can be displayed on the screen after being touched by a light pen. A flare data base for a time period can be constructed interactively in a relatively short time by a user with an archive of appropriate solar images.

3.2 Comparing different data or information.

The use of a computer graphics and image processing system in comparing different data sets can be crucial to the discovery of hitherto unrecognized correlations. Data sets of different modalities (such as different spectral lines, or magnetic field distribution versus x-ray brightness), or of the same modality but taken at different times (say a solar rotation apart), can be visually compared by splitting the display screen and showing the different phenomena side by side, or by overlaying contours onto images, or by differencing images and highlighting changes in color, and so on.

3.3 Transforming data to different coordinate systems.

Because of the sun's rotation, a solar feature is continually observed from different perspectives as it crosses the visible solar disk. To understand how a feature changes with time, it is more useful to observe the feature from the same viewpoint during that period. With a computer graphics capability, well-defined boundaries in the photosphere or chromosphere can be outlined interactively on a digitized solar image, and if the central meridian and solar inclination of the image are known, a new simulated solar image can be displayed showing the feature viewed from any perspective. Thus a solar feature observed over several days as it crosses the visible solar disk can be redisplayed in heliocentric coordinates to reveal intrinsic time changes.

When periods of cloudy days interrupt observations, or when a feature reappears from the farside of the sun, it may be desirable to get a rough idea of how the feature developed during the time it was not observed. One possible approach is to use the computer to interpolate the feature contours or surfaces between the instants of observation. The procedure is similar to that used in cartoon animation by computer, where the artist draws only several time steps, and the computer creates the intermediate images. Such procedures may be of particular use in following coronal hole boundaries and other large-scale synoptic features to understand the underlying dynamics. Perception of dynamical processes is usually easiest when the observed time flow is uniform and cinematic studies are possible.

3.4 Creating and interrogating data bases.

In my opinion, the most important potential uses of computer graphics in solar physics relate to synoptic data bases, in particular, to (1) the creation, maintenance, and continual updating of new kinds of partially-processed long-period synoptic data bases, (2) the ability to access, process, and integrate synoptic data quickly, to the point where high resolution and global information can be studied together as part of the same scenario, (3) the ability to compare different synoptic data bases in the same format, for example as spherical or planar maps, (4) the ability to view, portray, or overlay synoptic information in theoretical models of solar phenomena.

The purpose of science is to ask questions of nature and to seek answers from the observational data available. Collections of raw solar images can be collected and archived on video tape, but unless specific data can be retrieved on demand and interrogated, they are not of much use for scientific research. The number of solar images collected each year in each solar observatory is enormous. At present, those images are usually examined at the time they are taken for any features of current interest to the observer. Archiving is usually up to the individual observer. The retrospective examination of solar images taken by another observatory is usually painfully tedious, most often requiring the interested scientist to travel a large distance and to impose on the time and goodwill of his colleagues.

Usually what a theoretical scientist needs to know of past solar activities involves synoptic types of information. What was the distribution of the line-of-sight magnetic field component in the photosphere during a certain period? What was the hydrogen-alpha intensity distribution in the chromosphere? What was the coronal density distribution? Where were the prominences and filaments? What was the x-ray activity? Where were the coronal holes? and so on. It is my contention that much of this kind of information can be put into data bases, and that ten to twenty such data bases for a 22 yr period can be stored on a single removable disk, on tape, or on read-only optical memory, and mailed to every observatory. Maintenance of such synoptic data bases would provide every solar researcher access to data not available in-house, and in a form that does not require processing of raw images taken with unfamiliar instruments in unfamiliar spectral regions.

To set up a data base, an observatory must decide on a format for covering the solar surface or coronal volume, digitize the images or data into pixels, average the measurements in each pixel, and process the data for noise or for extracting the meaningful information. Measurements should also be averaged in each cell for several days as the region crosses central meridian or the limb of the sun (as the case may be). The averaging of the data removes noise and short range and short period

fluctuations, at the expense of high spatial and temporal frequency information. High-frequency filtering helps to avoid problems of aliased (undersampled) data. A volumetric data base for coronal density, involving voxels, can be constructed from a time series of processed coronagraph images. Original images or data from which a data base is constructed must be carefully archived in case errors or unusual events are found, or in case a new data base is suggested by future discoveries.

As an example, raw line-of-sight magnetograph data is not readily usable (because of quantity and quality) by the average solar physicist. But if the observer were to first filter, correct, and average the data so as to generate an equal-area cylindrical projection of say 360 equal longitudes by 180 equal divisions of the solar diameter, then everyone would be able to recall the field information from a data base and use it for coronal field and large-scale solar activity studies. If we assume that the data base is generated from solar regions within 30 degrees of central meridian passage and is updated every 2 weeks for a 22 yr period, then the size of the data base is

$$(180 \times 360) \text{ pixels/map} \times 26 \text{ maps/yr} \times 22 \text{ yr} = 37 \times 10^6 \text{ words}$$

which can easily fit on a small disk. With optical read-only memory, 20 such synoptic data bases, from hydrogen-alpha to x-ray intensities, could be stored and mailed in a small package. Once such a data base package is placed in the computer, a user could display information from the same or different data bases or from the same or different time periods in any mapping projection and in overlay, split-screen, or multi-color to search for correlations not previously apparent.

Suppose a numerical solution of a set of partial differential equations purports to show the evolution of certain patterns in the global photospheric magnetic field. Then with the proper software, the entire 22 yr data base could be interrogated to search for these patterns in minutes or less.

4. Setting up a Graphics and Image-Processing Facility

There is little doubt that a graphics and image-processing facility would greatly enhance information processing and understanding in solar physics, and lead to new perceptions of photospheric, chromospheric, and coronal interactions in both small and large scales of space and time. Having said this, however, I wish to emphasize that the setting up of an image-processing/graphics/database capability such as described above is not an easy matter in practice. Although hardware exists, it is still quite expensive (presently \$100,000 to \$300,000) for adequate systems. Image processors and display equipment still allow only one or a few users at a time, so that duplicate hardware systems would be needed by each observatory

and research group. Software primitives for computer graphics are still non-standard, although efforts toward standardization are beginning. Interfacing of image-processing facilities with computers is not trivial; languages, operating systems, and data transfer must be compatible, smooth, complete, and simple to understand and use. Long-term reliable funding must be identified to create and maintain specific data bases. Different observatories must commit themselves to set up and maintain these data bases. Some solar physicists must be willing to sacrifice a considerable part of their research time to create rather complex applications software that enables other users to meaningfully interrogate the image and information data bases.

The creation of synoptic data bases ideally should be a cooperative international project, but because of the expense and limited availability of equipment, and the absence of standardization, it may be more efficacious for a relatively small number of regional observatories and research groups to begin the effort.

The interested groups must first agree on the overall vision of where they want to be a decade hence. What problems should be addressed and with what image and synoptic data bases? What techniques and applications software would be necessary to interrogate and display the data bases in a useful and suggestive way? What types of image processing and software capability would be required to reduce incoming raw data quickly and cheaply and then update each synoptic data base?

Once the overall vision is agreed upon, the task of software acquisition for applications, image processing, and graphics must be assessed. Here is what we want to do and here is what is available. How much of what is available can be purchased with the funds available? What is the cost in time and money of developing software in-house versus purchasing software versus providing grants to university departments?

Once a realistic appraisal is made of the problems to be attacked and the resources available or potentially available, the specific funding, personnel, institutions, and equipment must be identified. Commitments should be made for the purchase of identical equipment for each research site. Protocols for data transfer and communication, etc. must be agreed upon. If different groups have different host computers, then computer interfacing and software transportability problems must be examined carefully. My feeling is that in the long run it would be better for all groups to purchase the same or compatible (32 bit) superminicomputers as hosts and strictly enforce program, system, and data compatibility.

A few points may make good sense if a computer graphics specialist is not expected to reside at each research site:

Make sure the graphics/image-processor hardware comes with software that is well-documented and easy for the non-expert to use. A user should be able to boot the system, write graphics programs in a high-level computer language (for example, by calling FORTRAN subroutines or PASCAL procedures), compile, link, and debug rapidly and quickly, and in general feel comfortable in working with the system.

Make sure the graphics company maintains the software, and, if possible, that newer software developed by the company will be compatible with the system purchased and not require new hardware.

Get as much software and memory in the image processor as you can afford. In the long run, this will allow much faster processing of images, rotation of 3-D objects, Fourier analysis, hidden-line removal, etc. As much processing as possible should be done in the image-processing system rather than in the host, provided every research site has the same equipment. If the same hardware cannot be purchased by every site, then host-resident software routines should be written as modules in a standard high-level computer language, to be replaced by faster image-processor-resident routines (with the same calling sequence) by each research group as needed.

Use direct-memory access (DMA) between disk memory and image processing hardware, so that the host computer can do other processing. Transmission rates of 1.5 Mbytes/sec are state of the art. Optical memories and fiber-optic transmission lines may soon be orders of magnitude faster.

Beware of any system advertised as a single chip that can do everything! This usually means a bottleneck. The main bottlenecks occur during data transfer: host to host, host to image-processor, host to disk, image-processor to disk.

Don't purchase the first model (unless you get special funds or a special discount on the price)! Let the company and other users get the software and system bugs out. Today's recommended technology may be somewhat slower than tomorrow's, but it is proven.

See if you can get a smart digitizer which lets you remove least significant bits (shift down) for noise reduction while the image is being digitized and sent to image memory.

5. Afterthoughts

Image processing and computer graphics are still frontier technologies. More and more realism is becoming apparent in computer graphics for motion picture entertainment, simulators for training pilots and machine operators, visualization of space missions, and city planning. Industries are beginning to rely on computer-aided design and computer-aided manufacturing (CAD/CAM). Image processing, in particular, image reconstruction, is now one of the most important components in diagnostic medicine. Robot vision systems will depend heavily on real-time image processing techniques, both hardware and software, still under development. In short, there is the great expectation that all aspects of human life will be profoundly affected when humans can communicate quickly and efficiently with computers by means of photographic and television images, diagrams, graphs, pictograms, and symbols.

I have no doubt that image processing and computer graphics will also have a profound effect on all fields of scientific research as the techniques become cheaper and more user-friendly. Solar physics is an ideal testing ground for applications of these techniques in science. There is always a backlog of images to be examined, numerous visible solar features to be studied and correlated in space and time, and large and small scale distributions of scalar, vector, and tensor fields to be studied dynamically in every time scale. Moreover, numerical techniques are widely used to test hypotheses on solar dynamos, flare initiation, particle acceleration, etc. All of this effort can be organized, archived, retrieved, displayed, compared, disseminated, correlated, and studied much more effectively with the techniques of computer graphics, image processing, and data base structures than by any other means. I strongly recommend that efforts be initiated to plan for the use of these techniques in solar physics.

Acknowledgments

I would like to thank R. Wallace of my department for helpful discussions on the state of the art in computer graphics and image processing.

Graphic Displays of Vector Magnetograph Data

D. M. Rabin and E. A. West, NASA Marshall Space Flight Center

Abstract. We summarize our experience with graphic displays that have proved useful in dealing with vector magnetograph data in three settings: real-time control, analysis, and final presentation. Among the topics discussed are: flexible, implicit data-scaling; geometrical transformations; methods of comparing fields (e.g., transverse vs. longitudinal; observed vs. computed; one time vs. another); displaying the magnitude and direction of the the transverse field; minimizing the display time of serial graphics devices; graphic file structure; and graphic interaction with operators and observers.

1. INTRODUCTION

Data involving a three-dimensional vector field require a greater variety of graphic displays, both for analysis and effective presentation, than do one-dimensional data such as line-of-sight magnetograms. Here, we first give summary recommendations for hardware and software appropriate to vector magnetograph data. Some of the recommendations are general, in the sense that they apply to many graphics applications, but are included because we found them particularly relevant to vector field data.

Although we discuss certain procedures, such as geometric transformations and implicit data scaling, that can greatly affect the apparent content of an image by altering characteristics of the display, we exclude many useful procedures that involve extensive arithmetic manipulation of the data prior to display; such image processing is outside the scope of our discussion. Also, we do not address problems associated with the display of high-resolution line profiles (the MSFC vector magnetograph uses a Zeiss birefringent filter as a spectral isolator).

2. SUMMARY RECOMMENDATIONS

Display Hardware

- * Direct Memory Access (DMA): consider it essential.
- * More screen resolution than you thought you needed: preferably 1024 x 1024.

- * Greytones (≥ 16) first; pretty colors second.
- * graphic memory with separately addressable bit-planes
- * cursor: easy manipulation and two-way communication
- * hardcopy capability: preferably both low quality (cheap and fast, for everyday use) and high quality (for reproduction)

Software

- * As far as possible, buy it: it has already been done and it is cheaper than your time.
- * Look gift dinosaurs in the mouth: some public-domain packages are conceptually outmoded and heavily hardware-dependent.
- * Start with an integrated 2D-3D, device-independent graphics package that conforms to a recognized standard (GKS-VDI or SIGGRAPH/CORE)
- * Interactive
- * Flexible geometrical transformations
 - * At least: flip, shrink, and squeeze (for non-square pixels); preferably, general affine transformations.
- * Implicit nonlinear data transformations (see below)
- * Direct-access file structure
- * Ability to make simple comparisons (side-by-side, contour overlay, blink, subtract, divide)
- * Eventually, Bells & Whistles (filtering, pattern recognition, automatic registration, etc.)

3. DISCUSSION OF EXAMPLES

It would be tedious to justify the summary recommendations individually. Instead, we illustrate them with observational data.

PLATE 1. This is a typical four-panel display used to monitor vector magnetograph observations as they are taken. Clockwise from upper left, the panels are: B_L , the signed value of the longitudinal (line-of-sight) magnetic field; I , the unpolarized intensity (showing sunspots); B_T , the strength of the transverse field; and B_{Tp} , the calculated strength of the transverse component of the potential field corresponding to the observed

longitudinal field. Such side-by-side display is the most rudimentary form of comparison. In order to create multi-panel displays simply and flexibly, the graphics software should have the capability to define multiple "windows" or "viewports": subsets of the graphic space that can independently be assigned user coordinates and positioned within the display area.

Plate 1 illustrates the point that Direct Memory Access (DMA) should be considered a necessity, particularly in view of its diminishing cost. Operating at 9600 baud, a serial graphics device requires about 15 minutes to display Plate 1 -- up to twice as long if the host computer is significantly loaded with other I/O. In contrast, the data for Plate 1 are usually obtained in 2-5 minutes. Real time control is not possible. The advantage of DMA is just as important in the data reduction phase. The psychological effect of waiting 5-15 minutes to display an image, often repeated several times in order to find an effective scaling of the data, inevitably causes the investigator to examine fewer images than he or she would like.

If a serial device must be used, there are a few tricks for speeding up the display modestly (factor $\lesssim 3$). For example, most devices can paint a single large area faster than many small ones with the same aggregate area. If, as sometimes happens, a large fraction of the image is occupied by a single color (see upper right panel of Plate 1), time can be saved by first constructing a histogram of the frequency of occurrence of the various colors. Then, before individual pixels are painted, the entire image area is "washed" with the most frequent color; other colors are overpainted as necessary. [More sophisticated variants can be devised. Also, the histogram of colors need not be determined for each image separately; a statistical average will usually suffice.] For another example, some serial devices will allow the color lookup table to be changed without repainting the image; e.g., all red pixels may be changed to green by a single command. In combination with the technique of implicit data scaling discussed below, this capability may enable the user to experiment with many scalings at a cost of only one display time.

PLATE 2. This vector magnetogram illustrates geometrical operations and techniques for displaying the transverse field. The image was constructed in order to compare it with a satellite ultraviolet spectroheliogram (displayed on the same system) and a ground-based H α filtergram. Each of the images had its own orientation, field of view, and pixel scale. It was most convenient to match the digital images to the photographic H α image. In the case of the magnetogram, this was accomplished by (1) extracting a submatrix from the original 128 x 128 image; (2) flipping the image left to right; (3) squeezing the pixels into rectangles, because the image scale is different in the N-S and E-W directions; (4) shrinking the image to match the H α image scale; and (5) rotating the image to bring celestial north to the top. These geometrical operations should be considered a minimal requirement for effectively comparing images; they are relatively easy to implement in a window-graphics environment. More sophisticated procedures, such as nonlinear distortion or automatic registration of different images,

extend into the realm of image processing. However, somewhat more is possible at the display level if the device has separable bit planes that can be independently "panned" (translated) by local command. For example, 8 planes may be divided into two images of 4 planes each. One image is then panned over the other while the eye searches for the best match.

Direct access organization of the image file significantly speeds up the common procedure of extracting a submatrix for display. One convenient choice is to make each row of the image a directly accessible record. Also, several images can be concatenated in a single file without increasing the access time. In the case of the vector magnetograph, which may produce up to eight images from a single observing sequence (longitudinal field strength, transverse field strength, total field strength, azimuthal direction, inclination to the line of sight, and three unpolarized intensities), concatenation controls the proliferation of files and facilitates comparison.

Plate 2 displays the azimuth of the transverse field as undirected line segments (the 180° ambiguity is inherent in the Zeeman effect). In this 50×50 image, the individual segments are just adequately defined. In the leading spot (to the right), there is a tendency to see "stripes" (blocks of nearly aligned segments) in what a higher-resolution display shows to be a nearly radial field. This emphasizes the point that, to display the field azimuth without "jaggies" or angular quantization, considerably more screen resolution is needed than to display a simple image. In practice, with the 640×480 resolution of the device (Tektronix 4027) used to create Plate 2, effective azimuth plots are limited to 50×50 .

Another factor that can restrict the size of the displayed image is the limit on the number of colors that the device can display simultaneously. This limit is eight for the Tektronix 4027. In order to obtain a useful number of displayed levels (say, 12-20), it is necessary to employ patterns of adjacent screen pixels that combine to simulate intermediate shades. With a minimal 2×2 pattern matrix, each image pixel requires at least four screen pixels, and the resolution is effectively halved.

A final point concerns the convention (suggested to us by R. Moore) used to display the strength of the transverse field: the bipolar scale shows $|B_T| * \text{sign}(B_L)$, the magnitude of the transverse field multiplied by the sign of the longitudinal field. This convention highlights the familiar neutral lines of the longitudinal field and facilitates comparison with longitudinal magnetograms from other observatories. Thus, it is easy to see that the neutral line on the left is highly sheared along most of its length (the azimuth segments are nearly aligned with the neutral line), indicative of a highly nonpotential field. On the right, in contrast, the neutral line around the leader sunspot shows little shear (the azimuth segments are roughly perpendicular to the neutral line), indicative of a local field configuration that is close to potential.

PLATE 3. These spectroheliograms, taken with the Ultraviolet Spectrometer and Polarimeter (UVSP) aboard the Solar Maximum Mission spacecraft,

illustrate the utility of implicit nonlinear data transformations. The superposed contours of ultraviolet continuum intensity (showing sunspots) are derived from a separate UVSP image.

The aim here was to uncover the similarities and differences between the spatial structure of images in two lines arising from the lower transition region, NV ($T \approx 2 \times 10^5$ K) and $L\alpha$ ($T \approx 2 \times 10^4$ K). With natural, "well-tempered" scaling -- i.e., 12 equal subdivisions between the minimum and maximum of the data -- the nearly simultaneous images in the two lines look quite different (Plate 3a,b). After transformation (Plate 3c,d), a close correspondence between the two images is apparent; what differences remain are highlighted for further examination. Note that the two transformations are not the same and were not guessed a priori: they are the result of interactive experimentation. Another useful feature of the nonlinear scalings chosen is that they reveal detail in both the intense active areas and in the much weaker peripheries.

Similar results may sometimes be obtained more mechanically by the technique of histogram equalization: assigning, for each image separately, the same number of pixels to each displayed color. However, this requires sorting of the data and strays into the realm of image processing.

By an implicit transformation we mean systematic alteration of the breakpoints of the color table ("contour levels") as opposed to explicit arithmetic operations on the data. We may think of the set of breakpoints as a "data comb" with teeth that may be added, subtracted, separated, or compressed to accent different aspects of the image.

There are several advantages to using implicit transformations. The original data values are always preserved, no matter how nonlinear the transformation. This is desirable because the units of solar data usually have physical significance (gauss, number of photons, etc.) that is lost if the data are rescaled; in some systems, for example, the range of the original or transformed data is always mapped onto the range 0-255.

With implicit transformation, arithmetic is carried out only on the set of breakpoints, typically twenty or fewer, whereas explicit transformation of the data typically involves one or more floating point operations on 10^3 - 10^5 numbers, which may introduce a noticeable delay on a heavily loaded minicomputer. Yet the final result -- which pixels are assigned which color -- will be the same in both cases.

To consider one class of implicit transformations, let the highest and lowest break values be x_{\max} and x_{\min} (often, but not necessarily, the minimum and maximum of the data). Assume for the moment that $x_{\min} > 0$. Define transformed limits as

$$y_{\min} = f(x_{\min}), \quad y_{\max} = f(x_{\max}), \quad (1)$$

where f is a monotone function, e.g., $f(x) = \sqrt{x}$ or $f(x) = \ln x$. To obtain n

colors, choose $n+1$ transformed breakpoints spread between y_{\min} and y_{\max} according to

$$y_i = y_{\min} + \left(\frac{i-1}{n}\right)^p (y_{\max} - y_{\min}), \quad i = 1, n+1, \quad (2)$$

where $p > 0$. A value $p > 1$ progressively compresses the breakpoints toward the minimum, whereas $p < 1$ crowds breakpoints toward the maximum.

To obtain breakpoints for the implicitly transformed data, invert the transformation of Equation (1):

$$x_i' = f^{-1}(y_i), \quad i = 1, n+1.$$

For further flexibility, transformations may be composed: e.g.,

$$y_{\min} = g(f(x_{\min})), \quad y_{\max} = g(f(x_{\max})).$$

Then, after defining y_i as before,

$$x_i' = f^{-1}(g^{-1}(y_i)), \quad i = 1, n+1.$$

If the data are bipolar (like magnetograms or dopplergrams), it is usually desirable to apply the transformations separately to the positive data and the absolute values of the negative data.

Two other useful devices may be mentioned: floors and ceilings, and a floating neutral point for bipolar data.

The presence of a floor means that all data smaller the floor value are "swept under the rug", into the lowest bin; data above the ceiling go "into the attic" (highest bin); and implicit transformation is applied only to data between floor and ceiling. This is useful for excluding a few "hot" or "cold" pixels and for enhancing discrimination and contrast in specific subregions (plage, cell centers, etc.). Two floor/ceiling pairs are maintained, one for positive and one for negative data. In this way it is easy to obtain a common scale on both sides of the neutral point even when the positive and negative data have different absolute ranges and after transformations have been applied.

A floating neutral point is just a simple case of implicit transformation of bipolar data, corresponding to addition or subtraction. It is useful when the physically significant zero point does not have a data value of zero. For example, the pixels in a dopplergram corresponding to

gas motionless with respect to the solar surface may have a nonzero value because of spacecraft orbital motion or instrumental effects.

Plate 3 illustrates pure "power" transformations: i.e., $f(x) = x$ in Equation (1) and $p \neq 1$ in Equation (2). Power transformations have the attractive property that the ratios of the resultant subdivisions are invariant to an overall change in the scale of the data. In Plate 3c, a floor is set at 7 and a ceiling at 187, followed by a power transformation with $p = 1.2$. Thus,

$$y_i = 7 + \left(\frac{i-1}{10}\right)^{1.2} (187 - 7), \quad i=1,11.$$

In Plate 3d, the floor and ceiling are at 1 and 70, and $p = 2.1$.

The bipolar scale of Plate 2 reflects simple square root transformations ($f(x) = \sqrt{x}$, $p = 1$) with a floor of 150 gauss (approximately the threshold of detection). Thus, for the positive scale,

$$y_i = \left[\sqrt{150} + \left(\frac{i-1}{6}\right) (\sqrt{1170} - \sqrt{150}) \right]^2, \quad i = 1,7.$$

Finally, as mentioned above, implicit transformation may allow the data to be rescaled without being redisplayed, an important consideration for serial devices. This will be possible if, first, the device allows global reassignment of colors or patterns ("change all red to green") and, second, allows a fairly large number (≥ 64) of colors or patterns to be present on screen at the same time. We start with a fine-tooth data comb: one that has many more teeth than the number of breakpoints (color divisions) ultimately required. Initially, there will be several (≥ 5) teeth between each true breakpoint. Dynamic rescaling is achieved by redistributing the colors among the teeth. Many teeth are necessary in order that, after a nonlinear transformation of the initial breakpoints, there is a high probability that there will still be a tooth near each transformed breakpoint. The process is illustrated for the square-root transformation in Figure 1.

PLATE 4. These maps of the azimuth of the transverse field illustrate another simple form of comparison using separable bit planes. Plates 4a and 4b are azimuth maps of the same region. Plate 4a (red) shows the measured azimuth map. Plate 4b (green) shows the calculated azimuth map of the potential field corresponding to the observed line-of-sight field. Since the two maps are on different bit planes, the operator can blink rapidly between the two using a programmed key on the display keyboard. Another keystroke displays both planes at once, the result (Plate 4c) being a red-green-blue sum of the separate maps that highlights areas of agreement (yellow) and areas in which the observed and the potential-field azimuths differ significantly (red and green).

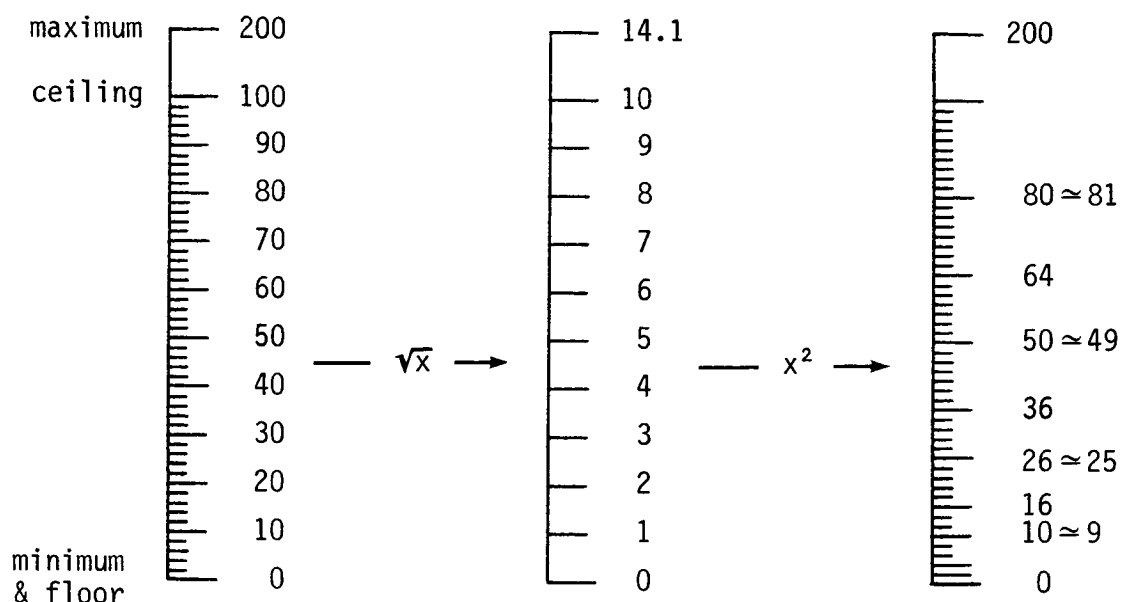


Figure 1. Implicit square-root transformation of a fine-tooth data comb. Any of the 100 data subdivisions between floor and ceiling may be assigned any of 10 displayable colors, before or after the data have been written to the screen; but the subdivisions themselves may not be moved (otherwise re-display would be necessary). Left, initially the colors are distributed evenly among the teeth. Color breakpoints are indicated by major (long) teeth -- e.g., all data in the range 50-59 are "green". Center, after forward transformation, new major teeth are chosen to be evenly spaced between floor and ceiling. Right, after inverse transformation, values that correspond to the major teeth in the center panel fall on or near teeth from the original comb, which are then adopted as new color breakpoints.

PLATE 5. Line graphics, such as histograms and scatterplots, should be considered an integral part of the image display system. If line graphics and image graphics are produced by the same software package, diagnostic plots can be intermixed freely with the images to which they apply. If separable bit planes are also available, plots can be superposed on their corresponding images and then removed when the analysis is complete.

The scatterplots shown are used to check for instrumental crosstalk (leakage) between circular and linear polarization, which is apparent as a correlation between the values of the linear Stokes parameter Q or U (plotted on the vertical axis, magnified four times) and the circular parameter V (plotted on the horizontal axis). Individual points are red or blue according to whether the datum came from the red or the blue wing of the $\text{FeI } \lambda 5250$ line. The sloping straight lines are fitted to the data and represent the best linear approximation to the crosstalk. Since these lines occupy a separate bit plane, the fits can be repeated using different subsets of the data without replotting the individual points. [Note that points near the origin have not been plotted in order to reduce the display time.]

The histograms shown are the distributions of values of various Stokes parameters $[U, Q, V \text{ and } (U^2 + Q^2)^{1/2}]$ over the image. They are used to identify small instrumental offsets introduced by a variable-tilt mirror (part of the correlation tracker) in the optical train.

A final comment concerns the interaction of graphics routines with the user. To be most useful to a scientist, image display and analysis routines should be highly flexible and interactive. On the other hand, an observer may not have the knowledge or the opportunity to respond to questions or select from long menus. One way to accommodate the needs of both types of users is to design highly interactive software that can receive input from more than one source. The scientist can make choices and supply information from the keyboard. For the observer, answers to most queries can be placed in a file tailored to a particular observing program; only a few questions need be directed to the operator's terminal. In this way, a single software package can serve real-time monitoring, quick-look data reduction, and detailed analysis.

The MSFC Vector Magnetograph program is supported by the NASA Office of Solar and Heliospheric Physics.

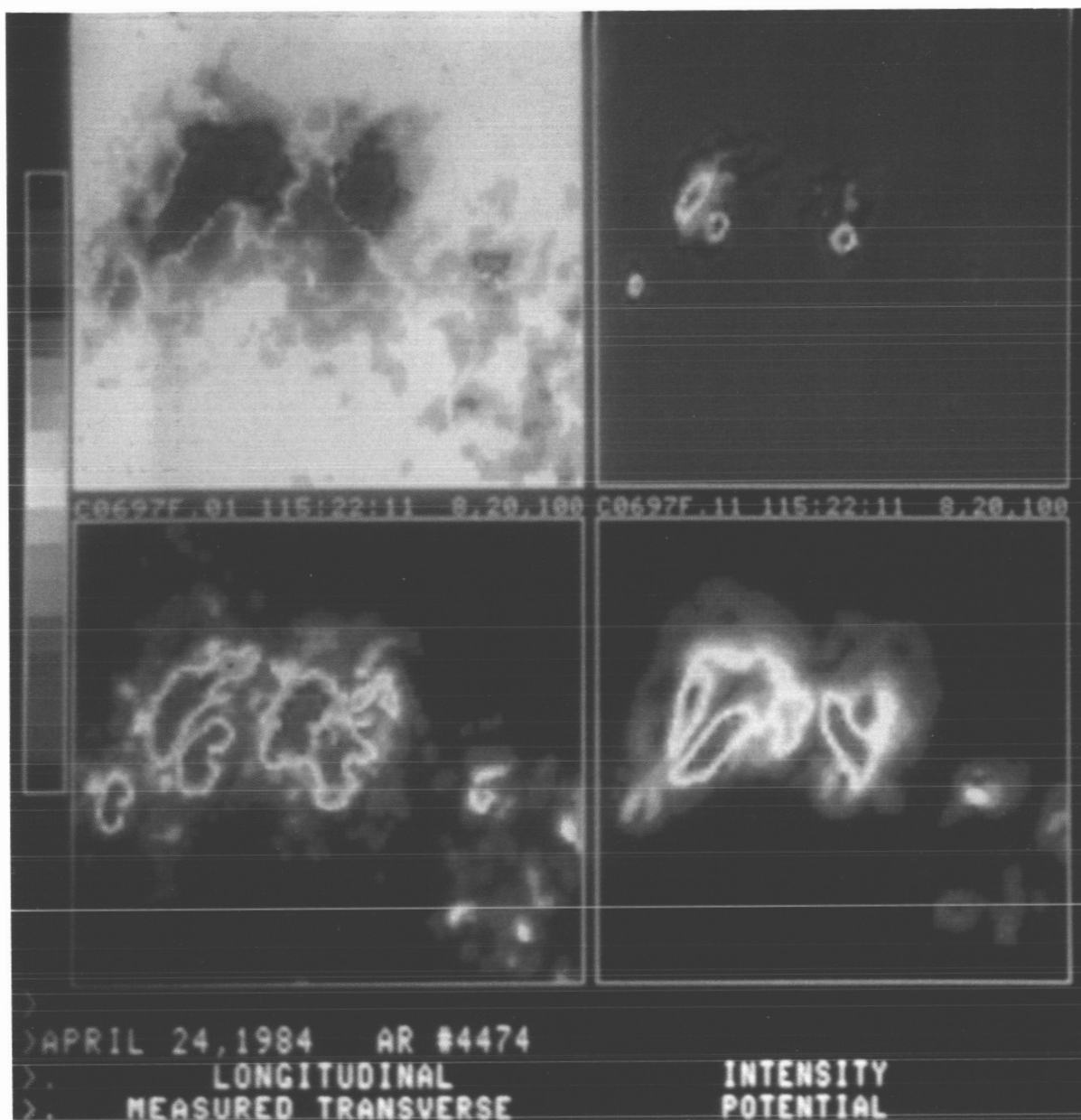


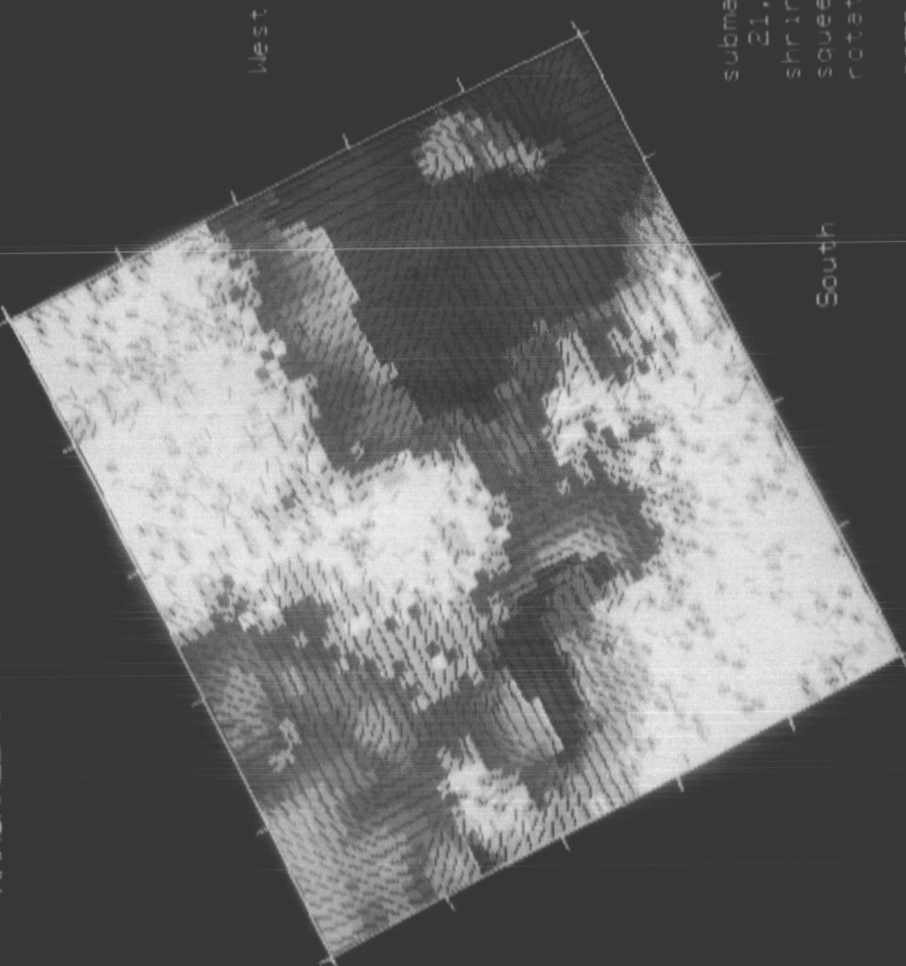
PLATE 1 (RABIN and WEST). Four-panel display used to monitor vector magnetograph observations in real time. Clockwise from upper left: B_L , signed value of the line-of-sight magnetic field; I , unpolarized intensity (showing sunspots); B_{Tp} , calculated strength of the transverse component of the potential field corresponding to the observed line-of-sight field; B_T , observed strength of the transverse field.

MARSHALL SPACE FLIGHT CENTER SOLAR SCIENCE

Vector
Magnetograph

Gauss
1170
990
721
538
382
252
150
0
-4
-150
-249
-375
-525
-701
-903
-1130

NO DATA



submatrix
21,16,50,50
shrink 0.85
squeeze 1.1
rotate 25

Active Region 2372

6 April 1980 97:19:08

Image: magnitude(transverse field) * sign(longitudinal field)
Segments: direction of the transverse field

PLATE 2 (RABIN and WEST). Vector magnetogram illustrating geometrical transformations (submatrix, shrink, squeeze, rotate), display of the transverse field azimuth (black line segments), and the convention of displaying the strength of the transverse component multiplied by the sign of the line-of-sight component. The scale at left, labelled in gauss, reflects an implicit square-root transformation (see text).

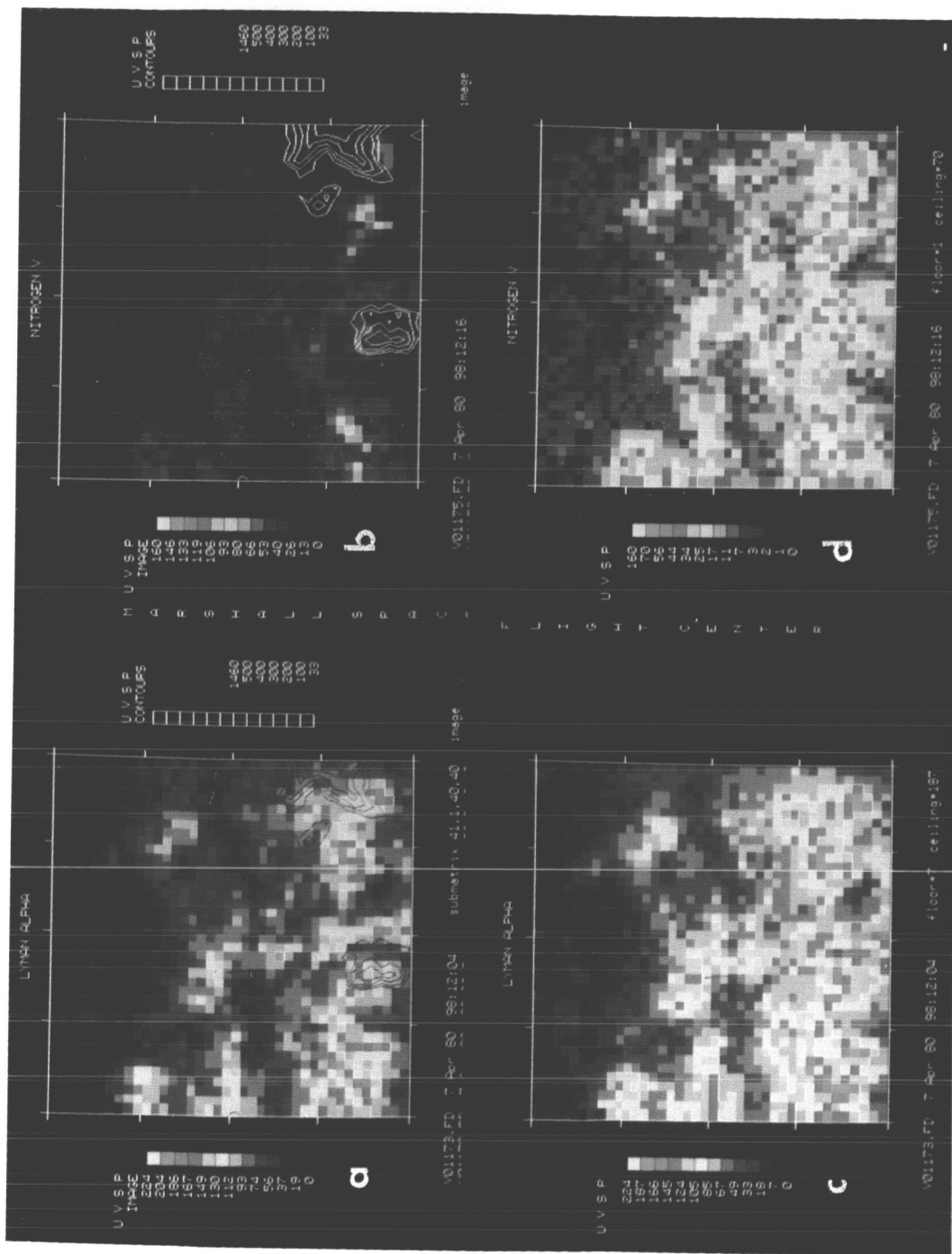


PLATE 3 (RABIN and WEST). Ultraviolet spectroheliograms (from the Ultraviolet Spectrometer and Polarimeter aboard the Solar Maximum Mission spacecraft) illustrating the usefulness of nonlinear scale transformations. (a) and (b): Lyman alpha ($T \approx 2 \times 10^4$ K) and N V ($T \approx 2 \times 10^5$ K), separated by 12 minutes, each linearly scaled; contours (showing sunspots) are from a UV continuum image; (c) and (d): the same data as in (a) and (b), but now with floors, ceilings, and implicit scale transformations chosen to bring out the close correspondence between the structure of the two images (the transformations are given in the text).

PLATE 4 (RABIN and WEST). Comparing two maps of the transverse field azimuth by superposing bit planes. (a) Measured azimuth map; (b) Calculated azimuth map for a potential field; (c) Superposition (red-green-blue sum) of (a) and (b): yellow indicates agreement, red and green indicate significant differences.

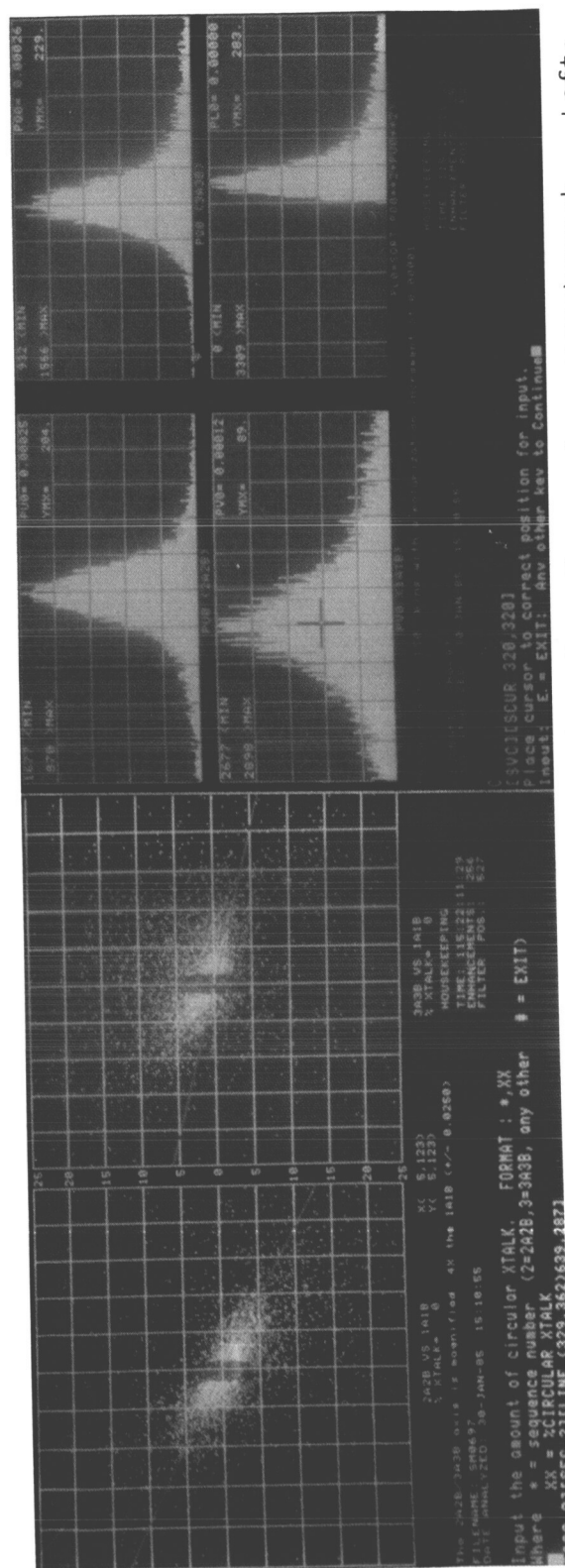


FIG. 21 (continued). Diagnostic line graphics for the vector magnetograph. Left: PLATE 5 (RABIN and WEST). Diagnostic line graphics for the vector magnetograph. Left: scatterplots, used to check for instrumental leakage between circular and linear polarizations. The fitted lines are used to correct for the leakage. Right: histograms of the values of various Stokes parameters, used to identify and correct small instrumental offsets.

LIST OF PARTICIPANTS

Dr. Martin D. Altschuler
Dept. of Radiation Therapy, Box 522
Hospital of the University of Pennsylvania
3400 Spruce St.
Philadelphia, PA 19104

Dr. Ulrich Anzer
Max-Planck Institute Fur Astrophysik
Karl-Schwarzschild Str. 1
8046 Garching
West Germany

Dr. R. Grant Athay
High Altitude Observatory
P.O. Box 3000
Boulder, CO 80307

Dr. A. Van Ballegooijen
Lockheed Research Laboratory
Dept. 52-13, Bldg 202
3251 Hanover Street
Palo Alto, CA 94304

J. J. Bao
Mechanical Engineering Department
University of Alabama, Huntsville
Huntsville, AL 35899

Dr. Richard S. Bogart
Institute for Plasma Research
Stanford University
Via Crespi
Stanford, CA 94305

Dr. Veronique Bommier
Observatoire De Meudon
Meudon 92190
France

Dr. Charles R. Chappell
Code ES51
NASA/MSFC
Huntsville, AL 35812

Dr. H. M. Chang
Mechanical Engineering Department
University of Alabama, Huntsville
Huntsville, AL 35899

Mr. N. P. Cumings
Code PF19
NASA/MSFC
Huntsville, AL 35812

Dr. E. Landi Degl'Innocenti
Istituto De Astronomia
Universita De Firenze
Largo Enrico Fermi, 5
50125 Firenze
Italy

Dr. A. C. DeLoach
Code JA01
NASA/MSFC
Huntsville, AL 35812

Dr. A. J. Dessler
Code ES01
NASA/MSFC
Huntsville, AL 35812

Dr. Audouin Dollfus
Observatoire De Paris
Section D'Astrophysique
92190 Meudon
France

Mr. James Dowdy
Code ES52
NASA/MSFC
Huntsville, AL 35812

Dr. Terry Forbes
Space Science Center
Demeritt Hall
University of New Hampshire
Durham, NH 03824

Dr. Allen Gary
Code ES52
NASA/MSFC
Huntsville, AL 35812

Dr. Mona J. Hagyard
Code ES52
NASA/MSFC
Huntsville, AL 35812

Dr. John W. Harvey
Kitt Peak National Observatory
P.O. Box 26732
Tucson, AZ 85726

Dr. William Henze, Jr.
Code ES52
NASA/MSFC
Huntsville, AL 35812

Dr. Ernest Hildner
Code ES52
NASA/MSFC
Huntsville, AL 35812

Dr. Satoshi Hinata
Physics Dept.
Auburn University
Auburn, AL 36849

Dr. J. L. Leroy
Observatoire Du Pic-Du-Midi
65200 Bagnères De Bigorre
France

Dr. Bruce W. Lites
Sacramento Peak Observatory
Sunspot, NM 88349

Dr. Timothy Loftin
Dept. Phys. and Astronomy
California State University, Northridge
Northridge, CA 91330

Dr. Boon Chye Low
High Altitude Observatory
P.O. Box 3000
Boulder, CO 80307

Dr. R. M. MacQueen
High Altitude Observatory
P.O. Box 3000
Boulder, CO 80307

Dr. Mitsugu Makita
Tokyo Astronomical Observatory
Mitaka, Tokyo, 181
Japan

Dr. D. L. Mickey
Institute for Astronomy
University of Hawaii
P.O. Box 209
Kula, HI 96790

Dr. R. L. Moore
Code ES52
NASA/MSFC
Huntsville, AL 35812

Dr. Yoshinari Nakagawa
Chiba Institute of Technology
Chiba 275
Japan

Dr. V. A. Osherovich
Space Environment Laboratory
NOAA/ERL
325 Broadway
Boulder, CO 80303

Dr. Eugene N. Parker
Lab for Astrophys & Spa Sci
University of Chicago
933 East 56th Street
Chicago, IL 60637

Dr. Alan P. Patterson
Big Bear Solar Observatory
Caltech
40386 N. Shore Drive
Big Bear City, CA 92314

Ms. Stephanie Patty
Mechanical Engineering Department
University of Alabama, Huntsville
Huntsville, AL 35899

Dr. Vic Pizzo
High Altitude Observatory
P.O. Box 3000
Boulder, CO 80307

Dr. Douglas Rabin
Code ES52
NASA/MSFC
Huntsville, AL 35812

Dr. David E. Rees
Dept. Applied Math
University of Sydney
Sydney NSW 2006
Australia

Mr. E. J. Reichmann
Code ES52
NASA/MSFC
Huntsville, AL 35812

Dr. E. Ribes-Nesme
L'Observatoire De Paris
DASOP
92190 Meudon
France

Dr. Paul H. Richter
Dept. of Physics & Astronomy
California State University/Northridge
Northridge, CA 91330

Dr. David M. Rust
Johns Hopkins University
Applied Physics Laboratory
Johns Hopkins Road
Laurel, MD 20810

Dr. Philip H. Scherrer
Institute for Plasma Research
Stanford University, Via Crespi
Stanford, CA 94304

Dr. M. Semel
Observatoire De Meudon
92190 Meudon
France

Dr. Raymond N. Smartt
Sacramento Peak Observatory
Sunspot, NM 88349

Mr. J. B. Smith, Jr.
Code ES52
NASA/MSFC
Huntsville, AL 35812

Dr. F. A. Speer
Code DS30
NASA/MSFC
Huntsville, AL 35812

Dr. D. Spicer
Code EZ-7
NASA Headquarters
Washington, DC 20546

Dr. Jan O. Stenflo
Institut Fur Astronomie
ETH - Zentrum
CH - 8092 Zurich
Switzerland

Dr. S. T. Suess
Code ES52
NASA/MSFC
Huntsville, AL 35812

Dr. E. Tandberg-Hanssen
Code ES01
NASA/MSFC
Huntsville, AL 35812

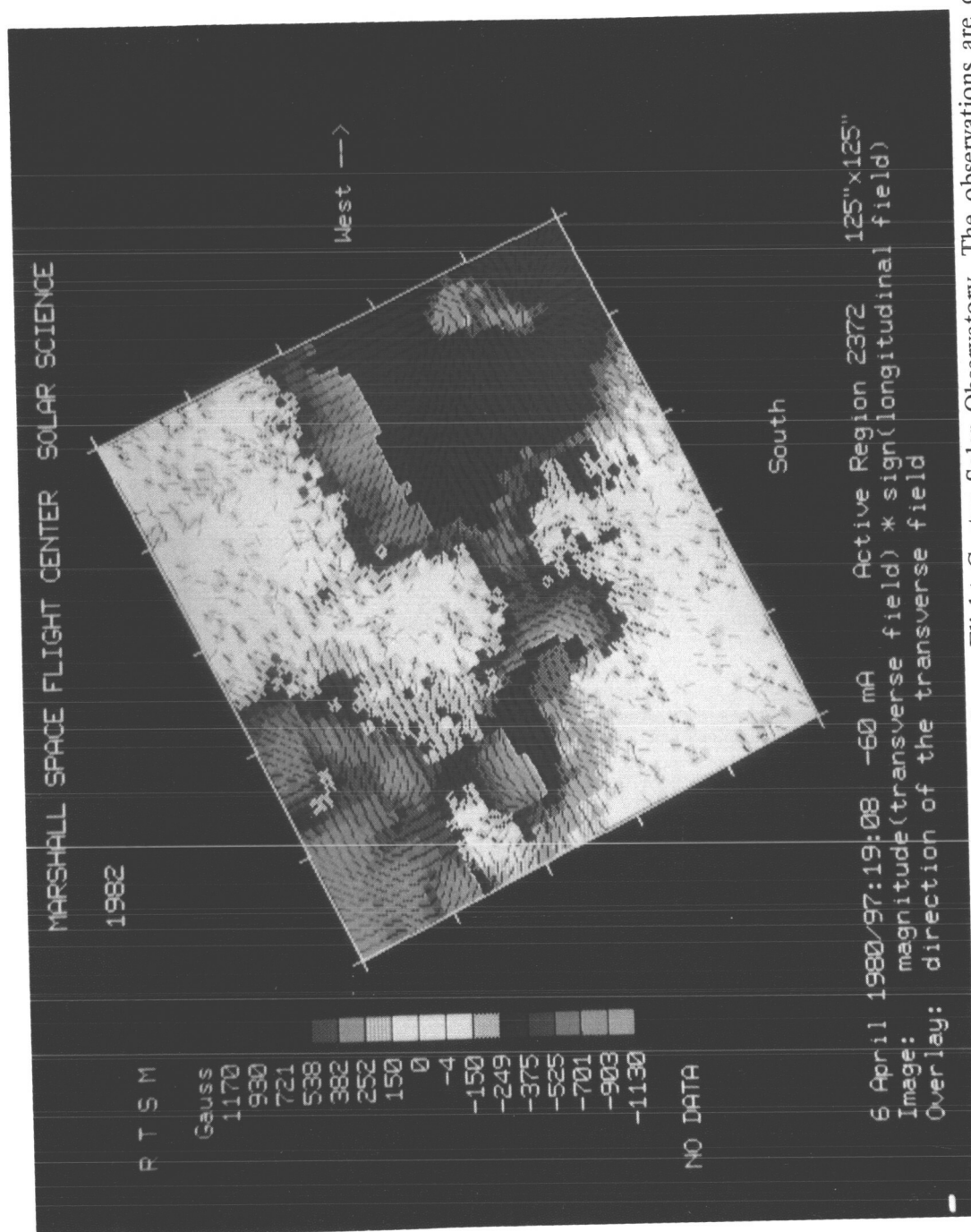
Mr. E. A. West
Code ES52
NASA/MSFC
Huntsville, AL 35812

Dr. Eberhard Wiehr
University Sternwarte
Geismarland Str. 11
D-3400 Goettingen
West Germany

Mr. Robert M. Wilson
Code ES52
NASA/MSFC
Huntsville, AL 35812

Dr. S. T. Wu
Mechanical Engineering Department
University of Alabama, Huntsville
Huntsville, AL 35899

Mr. Y. C. Xiao
Mechanical Engineering Department
University of Alabama, Huntsville
Huntsville, AL 35899



Vector magnetic field data from the Marshall Space Flight Center Solar Observatory. The observations are of Active Region 2372 on April 6, 1980. The image shows the locations of the line-of-sight field by colors: red (blue) intensities indicate positive (negative) polarities. The line segments show the orientation of the transverse field; the magnitude of the transverse component is proportional to the intensity levels in the color format. The area of highly nonpotential transverse fields along the neutral line in the southeast of this region was the site of frequent flaring during this time period.

1. REPORT NO. NASA CP-2374		2. GOVERNMENT ACCESSION NO.		3. RECIPIENT'S CATALOG NO.	
4. TITLE AND SUBTITLE Measurements of Solar Vector Magnetic Fields				5. REPORT DATE May 1985	
				6. PERFORMING ORGANIZATION CODE	
7. AUTHOR(S) Mona J. Hagyard, Editor				8. PERFORMING ORGANIZATION REPORT #	
9. PERFORMING ORGANIZATION NAME AND ADDRESS George C. Marshall Space Flight Center Marshall Space Flight Center, Alabama 35812				10. WORK UNIT NO. M-484	
				11. CONTRACT OR GRANT NO.	
12. SPONSORING AGENCY NAME AND ADDRESS National Aeronautics and Space Administration Washington, DC 20546				13. TYPE OF REPORT & PERIOD COVERED Conference Publication	
				14. SPONSORING AGENCY CODE	
15. SUPPLEMENTARY NOTES Sponsored by the National Aeronautics and Space Administration, University of Alabama in Huntsville, and National Oceanic and Atmospheric Administration					
16. ABSTRACT This report contains results of a conference held on May 15-18, 1984, at the Marshall Space Flight Center, Huntsville, Alabama. The content of the conference was organized into four sessions as follows: <ol style="list-style-type: none"> 1. Theoretical Understanding of Solar Vector Magnetic Fields 2. Techniques for Measurements of Vector Fields 3. Techniques for Interpretation of Observational Data 4. Techniques for Data Display. 					
17. KEY WORDS Solar Vector Magnetic Fields Magnetograph				18. DISTRIBUTION STATEMENT Unclassified - Unlimited Subject Category 92	
19. SECURITY CLASSIF. (of this report) Unclassified		20. SECURITY CLASSIF. (of this page) Unclassified		21. NO. OF PAGES 480	
				22. PRICE A21	